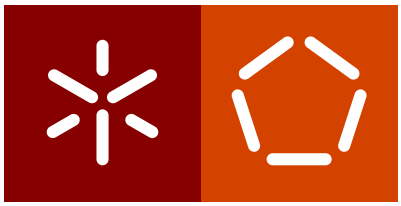
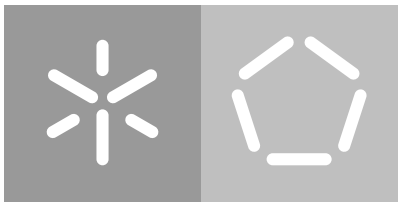


Universidade do Minho
Escola de Engenharia
Departamento de Informática

Pedro Miguel de Mata Rodrigues

Nutritional Management and Recommendations For Hospital Users and Medical Inpatients

January 2022



Universidade do Minho
Escola de Engenharia
Departamento de Informática

Pedro Miguel de Mata Rodrigues

Nutritional Management and Recommendations For Hospital Users and Medical Inpatients

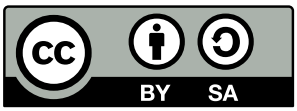
Dissertation Report
Mestrado em Engenharia Informática

Dissertation supervised by
José Manuel Machado
António Carlos da Silva Abelha

January 2022

DECLARATION

Name: Pedro Miguel de Mata Rodrigues
Civil ID Number: 15289478
Dissertation Title: Nutritional Management and Recommendations For Hospital Users and Medical Inpatients
Supervisor: José Manuel Machado
Co-Supervisor: António Carlos da Silva Abelha
Completion Year: 2021
Course: Mestrado em Engenharia Informática



Universidade do Minho, _____ / _____ / _____
Signature

_____

STATEMENT OF INTEGRITY

I declare to have followed the Code of Ethical Conduct of the University of Minho, and declare that this academic project was developed with integrity, and respecting the principles of ethics and morality, not having used plagiarism or any form of violation of third party rights, nor having falsified results throughout the process.

ABSTRACT

Nutrition is fundamental to human well-being and health, especially when applied to patients who need special health care. In these cases, it is crucial that each patient has adequate nutrition to meet their needs, in order to accelerate their recovery process.

Recommender systems make it possible to offer suggestions to users, adapted to their preferences and to previously obtained information about them. Food recommender systems are recommender systems applied to nutrition and diet. They are usually implemented feeding plans recommendation platforms based on food and the person using it.

In this sense, the existing gap in the use of these recommendation systems applied to nutrition in health care is notorious. This is mainly due to the difficulty in associating the nutritional value of each food with the needs of patients.

The main objective of this project is to fill the existing void, through the development and implementation of a platform that will allow the planning of meals taking into account the nutritional plan of the food and the specific needs associated with the users of the Vila Verde Social Canteen.

The use of machine learning algorithms will allow us to identify how the connection between food and patient requirements can be made, making this task possible, which is complex due to the wide domain associated with it.

This platform will be used for the generation of kitchen meal plans, which shall be produced using the algorithms developed after a bibliographic study and an investigation of the existing work, in order to understand how they can be implemented and which are the most adequate to the nutritional recommendations system.

KEYWORDS machine learning, meal plan, recommendation systems, nutrition systems

RESUMO

A nutrição é fundamental no bem-estar e na saúde do ser humano, principalmente quando aplicada a pacientes que necessitam de cuidados de saúde especiais. Nestes casos, é fulcral que cada paciente tenha uma nutrição adequada às suas necessidades, de forma a acelerar o seu processo de recuperação.

Os sistemas de recomendação permitem oferecer sugestões aos utilizadores, adequados às suas preferências e às informações previamente obtidas acerca dos mesmos. Os sistemas de recomendação de alimentos são sistemas de recomendação aplicados à nutrição e alimentação. Estes são usualmente implementados em plataformas de recomendações de receitas e planos de alimentação tendo como base a comida e a pessoa.

Neste sentido, é notória a falha atual no que diz respeito à utilização destes sistemas de recomendação aplicados à nutrição em cuidados de saúde. Isto deve-se maioritariamente à dificuldade na associação entre o valor nutricional de cada alimento e as necessidades dos pacientes.

Este projeto tem como principal objetivo preencher a lacuna existente, através do desenvolvimento e implementação de uma plataforma que irá permitir o planeamento de refeições tendo em conta o plano nutricional dos alimentos e as necessidades específicas associadas aos utentes da Cantina Social de Vila Verde.

A utilização de algoritmos de machine learning permitirá perceber como pode ser feita a conexão entre os alimentos e os requisitos dos pacientes, tornando possível esta tarefa, que é complexa devido ao largo domínio associado à mesma.

Esta plataforma será utilizada para a geração de planos de refeições da cozinha, sendo estes produzidos utilizando os algoritmos desenvolvidos após um estudo bibliográfico e uma investigação ao trabalho existente com o objetivo de perceber como poderão ser implementados e quais os mais adequados ao sistema de recomendações nutricional.

PALAVRAS-CHAVE machine learning, planeamento de refeições, sistemas de recomendação, sistemas de nutrição

CONTENTS

Part I

INTRODUCTORY MATERIAL

1

INTRODUCTION

The current document details the research methods and necessary analysis to be performed for the development of a platform for meal planning. This project was developed in the context of the dissertation project at the Master of Computer Science of Universidade do Minho.

In this first chapter, a brief contextualization about the project is presented, alongside with the primary goals to be accomplished with this project and the major motivation that led to its development. Closing this introductory chapter, an outline of the document's structure is presented.

1.1 CONTEXT

1.1.1 *Healthcare and Nutrition*

Nutrition is a key factor in everyone's quality of life and health conditions, and a even more important factor when applied to healthcare patients, because a good nutrition can be essential in their treatment and recovery process.

Medical Nutrition Therapy (MNT) is defined as the use of specific nutrition services to treat an illness, injury, or condition and involves two phases: assessment of the nutritional status of the client and treatment, which includes nutrition therapy, counseling, and the use of specialized nutrition supplements.

Besides user preferences in certain foods, health becomes more important as a factor in a food recommendation due to the increasing problems with unhealthy eating habits and their related diseases. [1, 2]

Researches and developments for a patient management system have traditionally been pursued within hospitals by specialized monitoring and decision-making support system. However, the recent increase of chronic diseases requires more ubiquitous healthcare systems. [3]

1.1.2 *Recommendation Systems*

Since the appearing of the first papers on collaborative filtering in the 90s, recommendation systems have seen an increment in alternative approaches and new ways to deal with the data that is available.

Over the last few years, the amount of data available is growing at an enormous rate, in relation to what was possible a few years ago. This data was discarded as it was only used for statistics and small experiments, but, as technology evolved and algorithms got more powerful, this data, when treated correctly, possesses and immeasurable amount of value. [4]

This is particularly true when talking about healthcare systems, because of the high risk and impact they can have on people whose life directly relies on the decisions the healthcare professionals have to take.

Although this systems can improve patients' quality of life by offering personalized suggestions, it's essential to keep in mind the potential complications of dealing with such practices. [5]

Recommendation systems use Artificial Intelligence (AI) methods to provide item recommendations to users. These methods are usually Machine Learning (ML) algorithms that classify previously obtained data and then compare it with the newly received to provide better suggestions. There's countless of use cases for these algorithms that range from robots to pattern recognition and self-driving cars. [6]

ML is a branch of AI that focus on how computer systems can learn from data and improve their accuracy over time without being programmed to do so. It simulates human learning and concede computers knowledge from real world, with the goal of improving performance based on the received information. [7, 8] These algorithms require an enormous amount of data in order to achieve a greater prediction efficiency, which means it's vital to get a lot of data related to the subject that's being analysed. This step is normally achieved through Data Mining. [9]

Data mining focus on extracting high-level knowledge from raw data. All this data holds valuable information, such as trends and patterns, which can improve decision making and increase success rates. [10, 11]

1.1.3 *Vila Verde Social Canteen and Kitchen*

Vila Verde Social Kitchen and Canteen (Cozinha e Cantina Social de Vila Verde) was built with the goal of serving meals to it's surrounding buildings, which represent Vila Verde's kindergarten, nursing home and hospital, as well as the underprivileged and homeless people of the city.

Hundreds of meals are produced every day at this kitchen and, depending on the day of the week, the number can vary based on many factors, depending on the day of the week, such as the number of patients currently in the hospital and how many people actually order a take-away.

The social kitchen provides the meals that are then served at the kindergarten, nursing home and hospital's own kitchen while serving those meals directly in the social canteen to families and individuals seeking financial and social support. This social canteen is also used by employees of the institution and other personnel.

1.2 MOTIVATION

Therefore, as the amount of meals produced at the social kitchen increased, it became a laborious task to manage what and how much to produce. The uncertainty related to the number of people that will benefit from social kitchen's support is also an important point to be considered, which increases the difficulty behind organizing this institution.

With this in mind, its directors looked for a way to ease the canteen's workflow to make sure everyone was served and no food was wasted. The development of a management platform that related patients and other canteen users' diet and diseases, together with Universidade do Minho, is an important step towards automating the management process.

The first step towards managing the kitchen while keeping in mind the patient's needs was to develop a web application to give guidance on the meal planning. The canteen already had a previously developed application, however it was old and plain basic, without major useful functionalities. It was mostly used to print weekly meal plans and to keep a record of previously cooked meals, and also to allow associating the different dishes produced daily with their correct section. This whole process is carried out manually by one of the people responsible for this area, and it is a complex and repetitive task, since it requires a deep analysis of the type of patients in each section and the nutrients that each patient needs daily.

1.3 DOCUMENT STRUCTURE

This document will be divided in 3 sections, each with several chapters:

- Introductory Material

  Chapter 1 - **Introduction**: Opening chapter presenting a contextualization of the work and outline the thesis motivation and objectives.

Chapter 2 - **State of the Art**: Presentation and analysis of some concepts necessary to proceed with this study and work previously done in the area, mainly related to machine learning and it's use cases in recommendation systems.

Chapter 3 - **Tools and Research Methodologies**: Analysis of the tools used in the course of the year and methodologies followed.

Chapter 4 - **Thesis Plan**: Description of the project's phases and outline of the development process.

- Data Preparation

  Chapter 1 - **Problem Understanding**: Introduction and analysis of the problem to be solved, with the goal of realizing what are the requirements to met on the development of this project.

  Chapter 2 - **Data Preparation**: Analysis of the data that will be required on the project, how it will be obtained and structured for later use.

- Solution Development and Conclusion

  Chapter 1 - **Solution Development**: Explanation of the decisions taken, system architecture and models and steps necessary to develop the solution.

  Chapter 2 - **Main Difficulties**: Essential problems found while developing the system and necessary steps taken in order to overcome them.

  Chapter 3 - **Conclusion and Future Work**: Description of the results obtained, discussion about what was achieved with the development of this solution and presentation of possible improvements and additions that can be made to the system with the intention of improving it.

2

STATE OF THE ART

2.1 INTRODUCTION

In this chapter, the discussion is focused primarily on the previously published articles and thesis work, what were the major difficulties found, and what needed to be done in order to reach certain goals. Starting with Machine Learning and going through what it is and how does it work, a small analysis of the various types of ML is performed, as well as how can each distinct case be applied into a practical and functional platform.

ML is implemented in several developed projects, integrated in different categories, therefore an examination was performed in order to understand what algorithm is normally applied to which domain and how the best performance can be obtained based on the examples considered.

Analysed projects will mainly fall on the recommendation system category, leading to a study of those projects to understand what their goal is and how it applies to the goal of this thesis. It is presented an explanation of how each algorithm was implemented and how it could be used for further improvement of this thesis. Based on what is usually the input received from the existent projects, either based on patient preferences or nutritional needs, it is considered how that data will influence the decision process and the output given to the user.

Concluding this chapter, sections studying work done using React and NodeJS are presented, in order to analyse and understand what are the best practices when developing a platform with high risk data like the one built for this thesis.

2.2 INTRODUCTION TO MACHINE LEARNING

Machine Learning is deeply connected with Artificial Intelligence and is focused on making systems learn from received data. These systems can understand how to perform tasks by generalizing previously executed examples, saving time and energy to what could be a time-consuming task when done manually by an human. [12]

It was first created from pattern recognition and grew into the amount of use cases that are available today. Its iterativity allows models to be independently adapted as they collect new data. There is more data available than ever in human history, and that was made possible by cheaper computational processing and more powerful and affordable data storage. [13]

2.2.1 *Machine Learning Approaches*

Supervised

Supervised learning focuses on making the computer learn a previously created classification system. Classification learning can be used for any problem where getting a classification is useful and easy to determine.

These algorithms use features and annotations in the training set and build a model with the goal of predicting the annotations of the instances in the testing set. Throughout this training, supervised ML algorithms can also allow the discovery of new patterns of classification, which give experts new results to analyse and new knowledge subject to various interpretations.

Essentially, if the task is to predict if an instance is related to a target variable, a set of instances known to be related to it can be used to build a model and use this model to get predictions for instances not used on training.

Supervised learning can be used with several learning methods such as classification and regression. [14] This method is commonly used on training neural networks and decision trees because these techniques depend highly on the information obtained by pre-classification. [15]

Unsupervised

Unsupervised ML techniques are executed without any given layers. The analysis is performed from raw datasets, meaning the understandings obtained from this data do not require expertise when working with these types of algorithms.

Most recent advances in unsupervised ML studies are due to the evolution of hierarchical learning, data clustering, latent variable models, dimensionality reduction and outlier detection. [16]

Semi-Supervised

Unsupervised learning is also often used in conjunction with supervised learning in semi-supervised learning. Processing the data before analysing it helps in the organization of the dataset, which leads to a better data structure, causing better feature representation and ease in finding patterns and structures in unlabeled data. [16]

Reinforcement Learning

Reinforcement Learning algorithms takes actions without knowing if they are correct or not, and learns through feedback received. The model is trained and tested at the same time and will adapt its future decisions based on what was considered right and wrong.

It is based on a reward system that benefits from good decisions and is punished when acting wrongfully. The algorithm then adaptes itself in order to maximize the expected reward. This method is mainly used on self-driving cars. [17]

Transduction

Transduction learning is a supervised ML algorithm where most data points of the test set are available for training without any labels.

These algorithms are mainly known to be instance-based, which means that problem instances are compared with previously trained instances. Most used algorithms using this paradigm are k-nearest neighbour and decision trees. [18]

2.2.2 *Supervised Machine Learning Tasks*

Classification

This supervised ML task is used to classify items in a data set into a predefined set of classes. A model is then built in order to predict this classes' attributes. The main goal of this task it so predict what is the target attribute for every item in the data set.

Classification tasks begin with a data set where the attributes of the elements are known. [19]

Regression

Regression task are in many ways similar to the classification task. However, in this case, the goal is to predict functions from input data into numerical or continuous variables.

2.2.3 *Unsupervised Machine Learning Tasks*

Hierarchical Learning

The understanding of simple and complex properties of input data, called features, from a hierarchy of multiple activations are what defines Hierarchical Learning. Features should also be informative, discriminative and independent. [16]

Data Clustering

Clustering's goal is to find patterns on unlabeled data in the form of clusters. This means that the input data is unlabeled and uncategorized and feeds the learning mechanism in order to obtain an output that is grouped in a meaningful way, by similarities, which will provide a better understanding about it's structure. [16]

Latent Variable Models

Latent variable are variables that are inferred from a set of algorithms and mathematical models from other observable variables.[20] A Latent Variable Model creates a relation between a set of these latent variables and the manifest variables. [16]

Dimensionality Reduction

Data can be represented in fewer or more dimensions, although real world data usually runs into the thousands of correlated dimensions. Dimensionality Reuction, in opposition to previously described tasks, focuses on the data points and theirs distributions in an N-dimensional space, instead of focusing on data lables. [16]

Outlier Detection

Outlier Detection excludes sample points that are distant from other samples in an axis representation. This task is normally implemented using the Nearest neighbor based and the cluster based methods.

2.3 Machine Learning on Nutrition and Healthcare

Machine Learning systems were able to make an impact in prediction and automation, by effectively learning and improving from input data. Combined with data mining algorithms, the amount of knowledge that can be obtained is immense.

As the amount of data increases, so does the knowledge that the machine has to learn, which means better predictions. [21]

2.4 Recommendation Techniques

Recommender Systems (RS) are software tools and techniques that can provide users with useful suggestions related to the type of wanted items. This is particularly used on movies, books and music programs, where the algorithm's fundamental goal is to give users the best suggestions for that type of item. [22]

RS are used for several reasons, such as the increase of number of items sold by a certain company, which can be achieved by suggesting to users items that are likely to suit the user's needs. This is beneficial for the user, considering the increased quality of the experience, and it is also beneficial for the company due to the higher gains.

These techniques are commonly powered by machine learning algorithms and data mining, that provide the data for these algorithms to properly work.

Multiple approaches are viable for developing recommender systems. The most commonly used ones will be described in this section. [22]

2.4.1 *Collaborative Filtering*

When using collaborative filtering, recommendations are made based on items that users with similar tastes previously liked. This is often referred to as "people-to-people correlation" due to the calculation being made taking into account user similarities. [22]

Suggested items are calculated based on items that were previously liked by users with similar tastes.

Collaborative Filtering is considered the most popular technique in Recommender Systems with Neighborhood-based and model-based being the most used methods.

Neighborhood-based methods focus on the relationship between users and items. On the other hand, model-based methods use Machine Learning and data mining to build predictive models. [22]

2.4.2 *Content-Based Filtering*

Content-based filtering, as opposed to collaborative filtering, bases recommendations on the item data, instead of the users data. This means that when a user is browsing for a certain item, the RS finds information related to that item in order to suggest other items with similar attributes. [8]

This approach analyzes a set of data related to an previously rated item and builds a model associated with the user's interests based on the features of those items. [22]

On contrary to collaborative Filtering, items they are not recommended because of the similarities between users. The recommendations are made based on other items. [8]

2.4.3 *Knowledge-Based*

There are, however, several types of items where the previously described methods are typically not suitable. This happens mainly on cases where the amount of ratings given by the user is not enough to provide useful suggestions.

Knowledge-based systems focus on resources not used by collaborative and content-based filtering approach, that can be achieved by "*exploiting explicit user requirements and deep knowledge about the underlying product domain for the calculation of recommendations*". [22]

These systems allow users to specify what data is useful to them, providing better control over the received results. This technique is, however, similar to content-based filtering in the sense that recommendations are based on the items' attributes, except that the knowledge-based technique does not delve into the relationship between users. [23]

2.5 Linear Programming

Conditional optimization models are mathematical models that aim to find the best solution to a given evaluation problem from a series of solutions. Each of these solutions is defined by a series of mathematical constraints that can be mathematical inequalities or equalities. [24]

It has several practical applications, and is also the basis for the development of combinatorial optimization.

An optimization problem consists of three essential points, namely a decision variable, whose value is unknown and is to be determined, an objective function that is to be optimized, and a number of constraints that must be obeyed by the decision variable. [25]

One of these models is called linear programming, and is used when the decision variables are continuous, when the objective function is a linear function, and also when the initial value of the constraint is a linear function.

A linear programming problem can have as an objective, as described above, minimizing or maximizing the objective function, while the constraints can have any linear combination. [24]

2.6 Food Recommendation Systems

Food recommendation systems provide users a recipes and dishes taking into account their needs and taste. This systems face a particular challenge in the Recommender Systems field, as a result of the necessary adaption to its requirements.

The amount of nutritional advice and the variety of nutrients and food flavors creates a difficult step that's crucial to overcome in order to get useful and correct suggestions.

This allied with the inherent risk of dealing with people's health increases the associated challenge.

The challenge associated with the development of food recommendation systems is expanded due to the difficulty in capturing user preferences and establishing a correct recommendation in order to improve user's eating behaviour.

Previously developed platforms and nutrition recommender systems were analyzed in order to perceive how each one obtains and treats user data and how suggestions are generated.

[2] have produced a nutritional recommender system that focuses on the use of food item databases that represent the nutritional value of each food. From these databases, micro and macro-nutrients were derived in order to produce a user profile. This method focuses mainly on ingredients previously consumed by the user, and creates recipes based on its nutritional profile and nutritional needs. This method does not consider user's health care, following only its intake history to produce recipe suggestions.

[23] have worked on a solution that directly aims obesity problems, focusing mainly on obese people. Their approach follows a knowledge-based method, producing its results based on physical data provided from the user, such as gender, age, height and weight, which are then used to calculate the body mass index. From this information, it's possible to perceive what the user's caloric needs are. Having this information collected, a user model is created aiming to create a context associated with it. A collaborative filtering approach is then used, making possible for the algorithm to understand what recipes and foods should be suggested to each used, taking into account data provided from users similar to the subject.

[26] have developed on a recommender system that is very similar to the one described bellow, where a user profile is first created in order to obtain suggestions from users in the same context. In this case, a questionnaire is also needed with the goal of obtaining user data, thus creating a user profile. This profile can be produced for healthy users as well as patients affected by chronic kidney disease, hypertension, or diabetes. Applying knowledge obtained from "nutraceutical characteristics of typical Calabrian foods", combined with users profile data, the algorithm analyzes whether a food item is appropriate for an user based on its health data, returning suggestions based on how accurate they are.

[27] have conceived a system that produces personalized nutrition recommendations based on the user information. When first using these systems, users need to complete a survey which will obtain their dietary intake and convert it into nutrients and calculate components associated with each food item. Recommendations are created based on nutrition experts data, that will initially provide nutrition recommendations manually, based on user data such as age, gender, weight and height and dietary intake. After several expert data is obtained, the process is then automated, granting users access to their recommendations. In order to avoid cold-start problems, this solution captures knowledge from national nutrition surveys, detecting correlations between dietary intake and socio-demographic data.

[28] have built a recommender system focusing on a hybrid recommender technique. This means both knowledge-based and content-based techniques are used. Knowledge-based technique makes use of users' data and food items in order to produce recommendations,

creating a relation between what items are adequate and meet user's necessities. The content-based technique produces suggestions based on food items features and ratings given by users, allowing the creation of correlations between food features (nutritional value) and user's previous selections.

[29] have developed a solution that aims to facilitate the user's task regarding to data entry. In this case, the user needs to answer an initial questionnaire, and then a series of food photos is presented in which he must identify those he likes. The image recognition system identifies the foods and creates a profile of the user based on the nutritional information associated with each food selected.

3

TOOLS AND RESEARCH METHODOLOGIES

3.1 INTRODUCTION

The development of a research project is constituted by a set of phases where every step is previously analyzed in order to understand which path should be taken, taking into account methodologies and technologies advantages.

For this dissertation project, the Design Science Research (DSR) investigation method was used in order to understand what solutions should be followed. Throughout every phase, several methodologies were studied and compared with the goal of distinguishing which was more appropriate for this project in particular.

In this chapter, a presentation of the DSR method is presented, explaining how each decision was made on knowledge creation based on the this method recommended approach. Furthermore, used tools are also described, examining the primary advantages over other possibilities.

3.2 DESIGN SCIENCE RESEARCH METHODOLOGY

Design Science Research methodology approaches a way of researching engineering and science related projects. DSR focuses on the development of artifacts with the goal of recognizing and improving the production of these artifacts.

Furthermore, DSR creates new knowledge through design and innovative artifacts, while using obtained knowledge to analyze and solve existing problems. Design Science Research usually *typically involves the creation of an artifact and/or design theory as a means to improve the current state of practice as well as existing research knowledge*. [30]

Moreover, this method comprises multiple steps, starting by the **identification** of the problem and the **definition** of a solution. In addition, the **design and development** of models and methods necessary to overcome them, preceding the **demonstration** of how these can solve the problem.2 This solution is afterwards **measured and evaluated**. [31]

This sequence is iterated, providing a better adaption to necessary changes until an ultimate solution is achieved. [30]

In order to meet all the requirements and solve current and emerging problems of Vila Verde Social Canteen professionals, the DSR methodology is used in this dissertation project.

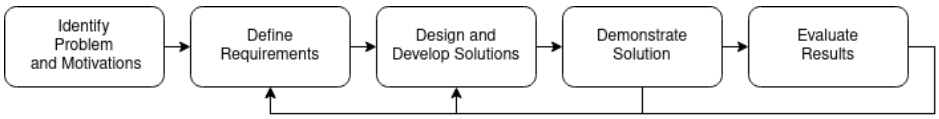


Figure 1: Design Science Research Methodology

3.3 TOOLS

A project of this kind consists of an integration between a set of tools and technologies, allowing the accomplishment of various tasks. Accordingly, the technologies used for the development of this project were: *Node.js*, an asynchronous event-driven Javascript (JS) runtime, which will serve as a web server; *React* and *React Native*, a JS library that allows the development of User Interface (UI) and an application framework that provides the ability to build native apps with the same code, respectively; *MySQL*, a relational Database Management System (DBMS), that will store users and other required data.

3.3.1 *Node.JS*

Node.js is a cross-platform runtime enrivonment for developing server-side Javascript applications. It is built on Chrome's JS runtime, which provides easily developed fast and scalable network applications. Being single-threaded could reduce it's scalability possibilities due to the being less prominent to execute concurrent functions but Node uses an event-driven, non-blocking I/O model, meaning it's lightweight and efficient, making it possible to use less resources for more requests. [32]

One of the most important features that Node provides are modules. Modules can be described as extensions for Node, which help speed up the development process. These modules can be core, third party and local modules. Core modules come installed by default on every Node installation. Third party modules can be installed and will be imported into a local directory, and can be installed using Node Package Manager (NPM). [32]

NPM allows the use of modules written by the Node community, as well as allows publishing self-created local modules for others. This tool eases module management in Node projects and provides a useful amount of previously developed code, meaning more efficient project development. [32]

One of Node most used modules is Express.js, a minimal and flexible web application framework. This module provides several features useful for web and mobile applications. [33] Express provides multiple interfaces that detach complex tasks from the developer, granting the ability to build faster and more maintainable web applications.

Moreover, the previously described tools culminate on the construction of a web service. This web service was built having in mind REpresentational State Transfer (REST) architecture. REST is a model that provides distributed interactions between multiple systems following a series of principals, also called RESTful Web Services.

REST premise is clear, using Uniform Resource Identifiers (URIs) to represent resources, while the access to those resources is used using the same methods normally used for HyperText Transfer Protocol (HTTP): GET, POST, PUT, DELETE, PATCH, etc... REST architecture is stateless, hence state is maintained by the resources. [34]

This methods can be used to operate and execute Create, Read, Update, Delete (CRUD) functions. When making a request to certain URIs, a response is expected, so for a simpler and more adequate use, considering the dominant language of this project being Javascript, responses are returned using Javascript Object Notation (JSON).

3.3.2 *ReactJS*

Also known as React or React.js, is an open-source Javascript framework from Facebook that aims to allow developers the possibility of building fast, simple and scalable User Interfaces. React's allows the production of stateful components, which are one of it's most useful features, alongside with the use of Javascript XML (JSX), Virtual Document Object Model (DOM) and Redux. [35]

JSX extends JS syntax, allowing the introduction of rendering logic alongside with UI logic, such as HyperText Markdown Language (HTML). This gives developers the possibility to render components directly with JS values. [36]

This logical execution is also made possible due to stateful components. That is, components are classes that implement a render method that returns what should be displayed to the User Interface. Furthermore, each component has it's own lifecycle methods and state, which is an object where data can be stored and read inside that such component.

Whenever any of the state object values are modified, the render method is automatically executed, automatically updating the browser's Document Object Model (DOM). This moves the complexity away from the developer and abstracts most of the repetitive tasks. [35]

React uses the Virtual DOM, an abstraction of the HTML DOM. This is a present in the *ReactDOM* module. [35]

In order to ease the process of dealing with larger applications, where state is unmaintainable, the Redux library is many times used alongside React applications. Redux consists

of the store as well as actions and reducers. Actions contain information about what should be updated in the application state, alongside with the data necessary to make those changes. Actions are dispatched in response to users interaction and are the triggers that then send new data into reducers. These define how the appplication's state will be changed in response to the actions. The store is a JS object with special functions that holds the application's global state. [37]

3.3.3 *MySQL*

MySQL is a Database Management System that allows the definition, creation and control of database accesses. MySQL is a relational database that stores data grouped by tables and uses Structured Query Language (SQL) for communication with the database. This query language provides powerful syntax for beginner users as well as experienced ones.

3.3.4 *Django*

Django is a web framework developed in Python, which has as main goal the fast development of web applications.

Using this framework, it is possible to develop secure and easy to maintain web applications, taking into account that this framework removes much of the existing complexity in implementing this type of systems. [38]

This framework has a series of extras that allow the developer to focus on other issues of development beyond the most common tasks, which are then in charge of Django, such as authentication, content management, RSS feeds, among others.

One of the points that seems to indicate why this framework is growing so fast is the fact that it allows developers to work in Python, which is also one of the fastest growing programming languages in recent years. This combined with the fact that all Django software is open-source, means that users can be aware of the system they are working on.[39]

4

THESIS PLAN

The ongoing dissertation is composed by six diverse phases, with different goals and distinct motivations. During project's delineation, several questions came up related to each of the phases described below.

Phase I sets the beginning of the project but it's also extended throughout the other phases, as additional information obtained from scientific articles and available literature will be used in order to improve document's accuracy. This initial phase focus mainly on analyzing previously developed Machine Learning algorithms and recommender systems, and its use cases when applied to nutrition and healthcare systems.

During Phase II, platform requirements were obtained together with the social canteen managers in order to better understand their current issues and prime necessities, as well as recognizing the biggest faults of their previous systems.

Phase II was also extended throughout Phase III, while pre-dissertation was written, requirements were analyzed in the purpose of bringing them closer to the needs presented.

After the writing of the pre-dissertation report, containing the bibliographical analysis and the initial objectives of the project, the development (Phase IV) will begin, preceding the test and production phases that will be carried out in Phase V. This dissertation will be concluded with the writing of the final report, the starting point of which is this report.

4.1 PHASE I: STATE OF THE ART

The development of a research project requires a scientific basis based on research articles and related projects, and as such an in-depth analysis of the work previously developed, is necessary. In this way it is possible to understand which production methods are the most appropriate for the type of system to be built, following the scientific knowledge obtained over time, both in books and in articles associated with the subject to be investigated.

In the case of this project, the studied articles are mostly related to machine learning and learning algorithms, since these are the most used in recommendation systems, such as the one intended to be developed. Besides the algorithms that may be implemented,

applications and projects related to the area in which it fits, namely nutrition and health care, were also studied.

4.2 Phase II: Requirements Analysis

After the initial investigation, a survey and analysis of requirements was carried out. To this end, meetings were held with the managers of the social canteen in order to understand what the objective for this project was and what the main needs were.

In this way, it was possible to see what the crucial flaws in their current system are and how it could be improved in order to achieve the intended goals.

4.3 Phase III: Pre-dissertation Report Writing

This phase involves the writing of the pre-dissertation report, which should describe what will be developed in the following phases, and also present the objectives of the project to be carried out. This report will serve as the basis for the final thesis.

4.4 Phase IV: Solution Development and Implementation

During this phase, the project will be developed, taking into account the tools designated for its production. This phase will count with the creation of the platform, and consequent implementation of machine learning algorithms in it.

At this stage, the analyses and studies previously developed should be applied to develop an application that meets the objectives initially outlined in an effective and efficient way.

4.5 Phase V: Solution Test and Integration

After completion of the project development, it should be tested for bugs and errors that could compromise the integrity of the system.

During this testing phase, the system will also be implemented and tested with real users, with the goal of finding new bugs and realizing changes necessary to the system's optimization in a real environment.

4.6 Phase VI: Dissertation Report Writing

The dissertation will be concluded with the documentation of the entire project carried out in the phases previously described, as well as the justification for all the options taken

throughout the process. The dissertation will be based on this report, with extra information being added as the project development progresses and new knowledge is obtained.

Part II

DATA PREPARATION

5

PROBLEM UNDERSTANDING

5.1 INTRODUCTION

The Vila Verde Social Canteen has a meal distribution service for the different divisions of this institution. Thus, there are a large number of people whose daily meals depend on their organization and preparation. In order to facilitate this process, the social canteen has an internal system through which each department can make reservations for its patients for each week, since the number of meals can vary due to the variation in the number of patients in each department.

Even so, the system currently in place is extremely old and needs daily monitoring by those responsible for each department, thus hindering the fluidity of the reservation process, which can be very costly in monetary terms and in organizational terms, due to being a laborious process and practically impossible to keep updated in real time.

Therefore, this project starts initially with the objective of being a support system for the social canteen, and its main function is to improve the existing reservation system, removing some of the work from the management of the social canteen in the organization of requests, but also with the intention of adding some extra modules that can make the use of the canteen more appealing, not only for the patients but also for the workers.

The most recent platform was already under development, although it was only in an initial phase, which means it was necessary to add some modules that would be needed to achieve the requirements initially presented by the social canteen.

In addition to the referred modules that would be necessary to manage the reservations made by each department, the system would also have to include a meal planning module in order to meet the nutritional requirements of each department, taking into account the needs that each valency would have depending on whether its patients are younger or older.

5.2 REQUIREMENTS

The initial requirements of the system are based on a series of proposals which, taking into account the system currently in place in the social canteen, should be a complement to the functionalities already existing, and to the system already previously developed.

Taking then as a basis the previously developed system, the essential functionalities to add to it would be the following:

- **Integration with the invoicing system**: the invoicing system currently used by the social canteen (Primavera) will allow invoices to be recorded directly from the web app, thus making it possible to automate the meal payment and importation process.
- **Management of facilities reservations**: reservations are currently made manually, with communication by email or messages between the managers of the facilities and the person in charge of the social canteen, making this process complicated for the person in charge of the canteen as they have to organise the messages they receive and register them somewhere externally. The process of booking meals should be done by each person in charge of each service automatically on the web page, without the need for intervention by the person in charge of the social canteen.
- **Patient evaluations on the website**: patients should be able to evaluate the meals they have consumed in the canteen, and in this way allow managers to get an idea of the general opinion of each patient. The current process consists of distributing a sheet of paper with a questionnaire to the patients, which is a time-consuming process for both the patient and the person in charge. The evaluations will be recorded in a mobile application to be developed in another project, and should then be viewable on the website.
- **Weekly meal plan**: in order to make the process of creating the weekly menu more practical, it should be possible to do it automatically, based on the patients.
- **Personalised meals**: the junction between this requirement and the previous one is necessary, since in order for the generation of meal plans to be possible, it is necessary to take into account the needs of each patient, and as such, it is necessary that the generation process of the weekly plan be done taking into account this requirement, otherwise it would be equally manual and time consuming, as it is nowadays.

6

DATA PREPARATION

6.1 INTRODUCTION

The first step in developing such a solution is obtaining the necessary data and processing it.

In this case, the relevant data was partly obtained directly from the social canteen, as they already had a system in place with user data.

However, the vast majority of the data that would be fundamental for this solution was not yet available, so it was necessary to generate a series of dummy values, which will also be used in the system testing phase.

In this chapter, the steps to be taken to obtain these data are presented, as well as the necessary handling of the data that will have to be done in order to store it in the desired quality.

6.2 DATA COLLECTION

The data used for the development of this project were mainly complaints and reviews made by patients about the meals taken, obtained directly through the institution with which the project was carried out, namely Santa Casa da Misericórdia de Vila Verde, as well as information regarding food products, having been collected from the National Institute of Health, which has a list of foods together with their components and properties, and is anually updated.

6.2.1 *Wards*

Data collection regarding the services in which this system would be implemented was made directly with the elements responsible for the social canteen. In This way, information was obtained regarding the people in charge of each wards, as well as the users present in each wards.

Obtaining data from the people in charge of the social canteen provides more accurate information and does not require additional processing.

6.2.2 *Users*

As well as the data on the services, the data on users and staff was obtained directly from the people in charge. Obtaining this information was, nevertheless, a simpler task since the data of each user was already in the system in use in the social canteen, and therefore it was only necessary to export these data to the new system.

The data obtained about the users corresponds only to personal information such as name, age, among others, and no information was obtained regarding the health status and/or diseases from which the users may suffer.

Although this extra data is not necessary for the first management modules of the system, it was necessary for the meal recommendation modules, which means that in this sense it was necessary to create dummy data.

6.2.3 *Food*

Unlike the data regarding users and skills, the process of obtaining the food was more laborious, and it was necessary to obtain this information through the website of the Instituto Nacional de Saúde (INSA), which has an API that contains information about a series of foods, as well as their properties.

This way, it was possible to obtain the information about the food, together with its composition, unit of measure, among other information that would be crucial in the construction of this system.

The API used returns the data in JSON format, and the information for each food is found in different files, and as such, a process of reverse engineering was necessary to obtain the data for all the necessary foods. The format of each JSON file is as follows:

```
{
    "matrixUnit": {
      "id": 20,
      "code": "W",
      "descriptor": "per 100g edible portion"
    },
    "id": 2021,
    "name": "Robalo cru",
    "nameEnglish": "Seabass, raw",
    "importSource": "IS871",
    "foodCategoryList": [
```

```
        {
          "id": 83,
          "version": {
            "id": 2,
            "name": "imported data",
            "versionpublic": false
          },
          "subCategories": false,
          "name": "Peixe",
          "categorySynonymList": [
            {
              "language": {
                "id": 184,
                "name": "English",
                "code": "en"
              },
              "id": 83,
              "term": "Fish (includes fish dishes)"
            }
          ],
          "parent": {
            "id": 58,
            "version": {
              "id": 2,
              "name": "imported data",
              "versionpublic": false
            },
            "subCategories": false,
            "name": "Pescado (Peixe, Moluscos e Crustaceos) e Derivados",
            "parent": null
          }
        }
      ],
      "componentValueList": [],
      "langualCodes": [],
      "foodpublic": true,
      "ediblePortion": "38%",
      "foodEx2": null,
      "specificGravity": null
  }
```

In order to obtain all the necessary foods, a small tool was developed that allows obtaining this data automatically by going through all the pages of the API. This way, the obtaining process becomes automatic, and taking into account that this food database is periodically updated, it is thus possible to obtain the food quickly whenever necessary, by simply updating the local system's database with the obtained data.

This tool developed using **Elixir** uses concurrency to be able to obtain the huge amount of data available in the API practically instantaneously, thus having the data updated permanently and without the need to overload the local system.

6.2.4 *Meals*

The dishes database was, like the users and services, imported directly from the system previously used by the social canteen, thus being possible to use real data, allowing the tests to be carried out on the system under construction to be as close to reality as possible.

Although it was still necessary to process the data received, this task became simpler, even though it was necessary to add extra data for the system to be functional, since the amount of existing data was small.

6.2.5 *Diseases / Recommended Plates*

Unlike the information described in the previous sections, there is no database or even an organised list of data that makes an association between diseases or health conditions and food.

It was therefore necessary to obtain all this information manually by reading many research articles in this area, in other words, understand what the experts say about recommendations for patients suffering from certain health conditions to eat.

The phase of obtaining this data was a lengthy process, due to the difficulty in finding a large number of articles on this topic, and also due to the need to understand which recommendations were given and the "strength" of these recommendations, as this is such a sensitive process, since an incorrect decision may endanger the health of a patient.

Even so, it was not possible to obtain complete dishes or meals associated with certain diseases, but only foods that are recommended for those same diseases, and then it is necessary to do further processing in order to transform the information obtained into what is required for the proper functioning of the system.

For this stage, and bearing in mind that the purpose of the module using these data will, at this stage, only be used as a proof of concept, the most common diseases currently in Portugal were selected, and those in which food can be an important factor in the user's

state of health. Articles were analysed mainly respecting hypertension ([40–71]), obesity ([72–88]), diabetes ([89–98]), and anemia ([99–103]).

The data obtained was saved in csv files, which were organized in order to facilitate their reading into the database or, if necessary, a preliminary treatment prior to their storage, represented using structure on Fig. 2.

food	Disease
nozes	obesidade
nozes	hipertensao
ovos	diabetes
ovos	obesidade
tomate	hipertensao
..	..

Figure 2: Diseases and respective recommended ingredients

6.3 DATA TRANSFORMATION AND PREPARATION

6.3.1 *Data Normalization*

The data normalization process consists in transforming the initial data so that they can be more easily integrated with the developed system, avoiding data duplication and formatting errors, thus allowing greater data integrity.

In this way, the data normalization process was carried out in order to transform the data that had different values but that symbolized the same thing, being the most evident example of this problem the case of existing values in plural and singular in different sources, despite representing the same thing.

This process was carried out manually for each word obtained, and the words were transformed to their singular.

6.3.2 *Food Nutritional Data*

Regarding the treatment of the food data obtained, the attributes that would not be necessary in this system were removed, mainly attributes that would be used for internal identification of the food in the API of the INSA.

Taking this into account, they were distributed among the Alergenico, Categoria_Ingrediente, Componente and Ingrediente tables, using the Unidade_Medida and Unidade_Porcao tables as normalization tables for the units of measure and portions used, as well as the Ingrediente_Alergenico and Ingrediente_Componente tables, which are used as auxiliary connection

tables between ingredient-allergen and ingredient-component, respectively, thus allowing a many-to-many relationship to be established between the various tables.

The attributes represented in the Alergénicos table are its name and its id. As for the Componente table, the attributes used are the component id, its designation, the code (which allows its identification in different systems) and the unit of measure used by this component (a foreign key of reference to the Unidade_Medida table is used). Regarding the Ingrediente table, its name, the source from which the INSA imported this ingredient, its edible part, the relative density, as well as the category of the ingredient and the serving unit are recorded, the latter two being references to the auxiliary tables created.

6.3.3 *User Data*

The users' data were introduced into the database in the User table, which would be differentiated from the users associated with the Staff through the definition of their user type.

As mentioned in the previous chapter, the data regarding the users' health status was not available during the development of this project, so a small dataset with dummy data for the users was developed.

6.3.4 *Diseases and Recommended Nutrients*

The data on recommended diseases and nutrients went through a normalisation process as described in the section above in order to convert all words into their respective singulars. Subsequently, all words were transformed that all characters became lowercase letters.

In addition, normalisation was also performed with respect to special characters, such as accented or language-specific characters (as is the case of the ç).

This process has made it possible to establish a greater integrity of the data obtained and to limit the existence of duplicate information.

Part III

SOLUTION DEVELOPMENT AND CONCLUSION

7

SOLUTION DEVELOPMENT

7.1 INTRODUCTION

In order to develop a solution that met the requirements presented, it was necessary to understand what the necessary architecture for the system would be, bearing in mind that the information would have to come from different systems with different objectives.

In addition, the information would have to be stored in a centralised way, so that the data would always be synchronised without needing a constant communication between systems.

In this chapter there is a brief explanation of the decisions taken, followed by an explanation of the system architecture, not only about why the system is structured in a certain way, but also about the technologies chosen for each section of the solution to develop.

In the following sections, a detailed explanation will be presented about the modules developed and how they will meet the proposed requirements.

The different algorithms to be implemented will also be explained, especially in the more delicate situations, where it is necessary to develop more complex algorithms, that come from the use of different technologies, after studying them.

To finish this chapter, the difficulties encountered during its development will be presented, as well as the steps taken in order to overcome them, concluding with a brief conclusion about what was perceived with this development as well as what can be done in the future in order to improve this solution.

7.2 DECISIONS TAKEN

The initial architecture of the project consists of the back-end composed of a Rest API developed in NodeJS (Javascript) and another in Django (Python), and the front-end composed of a web page and a mobile application.

The web page will serve as a dashboard, from which the people in charge of the social center will manage the entire system, which is composed of the reservations of the users

and services, as well as all the properties of the food inserted into the system, and where they can also insert new menus and even new foods.

This dashboard can be accessed by elements responsible for the social canteen, but also by other elements, and will have an authentication system, thus limiting the access that each user has to the different modules.

In the case of the REST APIs developed, it was initially produced the NodeJS server that connects to the MySQL database, where all the data to be used in this system is stored. This API provides the frontend with the users, the food and its nutritional information, menus, reservations and other details that may be needed for the system to work.

In addition, the need arose to develop an API in Django that allows the generation of weekly meal plans based on the nutritional needs of each user. With this focus, the development of a new API was necessary since the use of Machine Learning (ML) algorithms forced this to happen, considering that it would be much more complicated to develop this system using Javascript. This server is connected to the same MySQL database, thus allowing the data to be kept synchronized without the need to register it in more than one place simultaneously, which offers greater efficiency in terms of system stability.

7.3 SOLUTION ARCHITECTURE

In order to continue the development of this solution, the tools currently in use were analysed with the objective of understanding which would be the most suitable methodologies for the development of a platform that would be easy to use, easy to integrate with other systems - since there was a possibility of the need for integration between this platform and the canteen's invoicing systems - and also scalable, in case it was necessary to add new modules as the social canteen grew. As tecnologias utilizadas foram as seguintes:

- Frontend: **React**. The use of ReactJS for the development of the system's interface was the most obvious choice, taking into account that the system previously developed was already done using this framework, which means that it made all the logic to continue the system's development with what was already being used, especially considering the need to implement functionalities in a short period of time, therefore not being necessary to develop the whole system from scratch.

- Backend: **NodeJS**. As in the case of the choice of technology to be used in the frontend, in the backend the choice fell on what was already in use in the previously developed system. Even though it was anticipated that it would be necessary to change the way some of the themes were addressed in the existing solution, it would not be necessary to redesign the system from zero, simply by making the necessary changes so that the new functionalities would be in line with what had already been developed.

- Database: **MySQL**. Being this the most used database management service nowadays, and taking into account the previous knowledge of SQL, MySQL was used mainly due to the simplicity in which it can be integrated with the tools in use, in this case NodeJS.
- Recommendation API: **Django**. For the development of the recommendation API, the Django framework was used. Taking into account the knowledge previously obtained in Python, and as such the knowledge that there would be several libraries that could make the development of the recommendation algorithms a less complex process, it was decided that Django would be the most appropriate for this section of the system. Furthermore, it would be necessary to establish a separation between the remaining modules and this recommendation module, since it could apply more pressure on the machine where it would be executed, and as such it might have to be executed in a separate system.

The interface communicates directly with the REST web server, which was developed using NodeJS together with the **ExpressJS** framework, which allows a project to be created from scratch already having a base structure, thereby allowing faster progress in the initial phase of the project. In addition to this framework, the **PassportJS** middleware was also used to manage authentication, as well as the use of JWT (JSON Web Tokens) to maintain the security of the endpoints, and allow blocking access of each route only to authorized users.

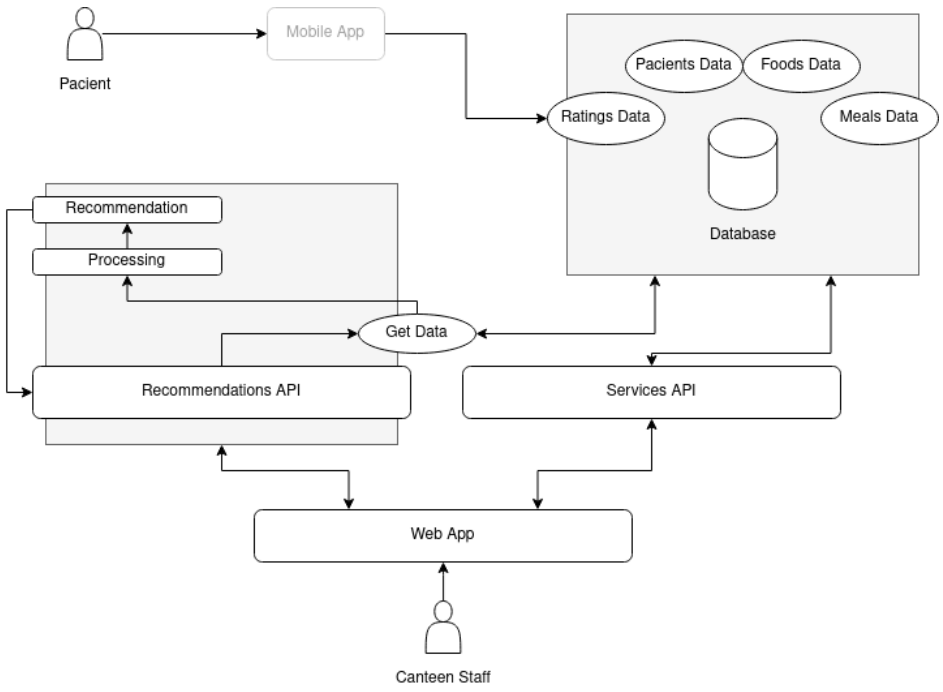


Figure 3: Solution Architecture

7.4 SOLUTION MODULES

For the development of the system, it was important to separate it into different modules, since each section of the project served different objectives, and in this way it would be desirable to obtain a better organization of the structure, also allowing different functionalities to be added as it was crucial to expand what had been already developed, unlike the existing project in the social canteen, which had been developed with one aim in mind, and thus made adding new functionalities quite complex.

Furthermore, the division of a system like this into different modules has numerous advantages, being simpler to manage, since each functionality is in its own module and it would be possible to activate or deactivate functionalities whenever necessary or whenever they compromise the structure in some way.

Due to the need to allow other programmers to take part in the project as soon as necessary, this system facilitates their integration, and it is even possible, if necessary, for each person to work on their respective functionality without compromising the code developed by the others.

This system also allows the reuse of code, or even its portability to other distinct projects, making it readily adaptable to the needs of each system.

For the development of the modules, the system requirements imposed by the social canteen were analysed profoundly and, from these, the following modules were created:

- Meal Management
- Meal Plan
- Section/Ward Management
- Reservation Management
- Meal Recommendation
- Weekly Meal Plan Generation

7.4.1 *Food Management and Meal Plans*

The basic module of this system is the management of meals that can be prepared in the social canteen. This module is made up of Ingredients, Components, Preparations and Dishes.

The components are the lowest element of this set, and can be represented in various units of measurement, such as grams, kilos or litres.

The ingredients have the various components associated with them, which indicate their nutritional data, such as the amount of calories, protein, carbohydrate, sugar, fibre, iron, among many others.

On top of these elements are the preparations. They make it easier for the members of the social canteen to enter the data so that there is no repetition of data. Preparations are made up of ingredients, and some examples of these are dry rice, fried potatoes or even fish fillets. A set of preparations constitutes a dish, so it is possible to create several dishes with variations of preparations.

The dishes have several characteristics that aim to improve their organisation within the system, such as the season to which the dish belongs, which can be summer or winter, with lighter and heavier dishes, respectively.

In addition, each dish can also be identified by its type, which can be breakfast, salad, dessert, main course or soup. Within these types, there are sub-types such as sultana soup, reinforced soup or normal soup, which allow the dish to be further customised to the user's needs.

The planning of weekly meals is, as described in the next section, a very time-consuming task since it is, in the system currently in force, done in a repetitive way in which it is necessary to manually enter all the dishes that will be made for all the meals of each day of the week. As such, a module was developed to associate each meal with the dishes previously registered, thus allowing this process to be simplified, since the dishes are created and it is only necessary to select them.

7.4.2 *Section/Ward Management*

Besides the weekly meal planning, it is necessary to divide the meals for each of the services, based on the users attending the canteen.

In this sense, a module was developed that intends to simplify this management, and it is possible to register users responsible for each service so that this process is done in a decentralised way, thus allowing that in the canteen it is only necessary to identify the needs of each service and plan the meals accordingly.

7.4.3 *Reservation Management*

Based on the requirements presented by those responsible for the project, it would be necessary to develop a module aimed at managing reservations, so that the staff of the social

canteen could have greater control regarding the quantity of meals to be prepared each day of the week.

As such, it would be necessary that the different responsible persons of each ward had the possibility to access the web page and register the quantity of users that would be available for lunch in each one of the days, thus establishing a reservation of meals for the users over which they have responsibility.

In this way, it would be possible to meet the need that they had initially presented, in which they showed their concern about the number of users being quite variable throughout the weeks, and as such a constant communication between the various wards and the central point of information (which would be the social canteen) would be necessary.

This module may later be expanded in order to allow reservations to be made directly by the users, allowing the work of those responsible for each service to be more efficient.

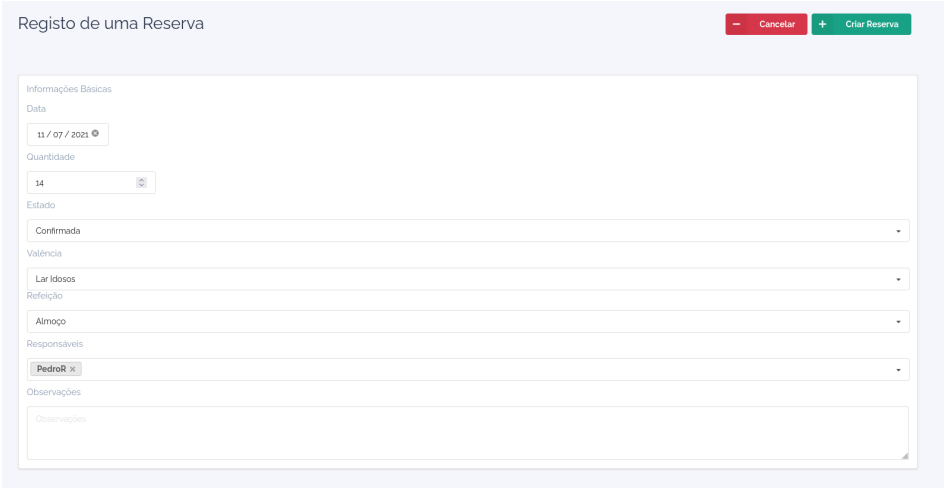


Figure 4: Creating a reservation as a ward staff

Reservations are managed on a weekly basis, but can still be booked up to a week in advance, allowing for better organisation by those wards that prefer it.

Figure 5: Reservations List

Reservations can have different statuses, such as ”Confirmed”, ”Cancelled” or ”Pending”. In this way it is possible to register reservations on a certain day of the week and then cancel them, either by the responsible member of the service that created it, or by a responsible member of the social canteen. This is important because cancellations or changes may occur

in various ways, whether because the booking was made too close to the time of the meal, or for reasons of lack of availability by the social canteen.

7.4.4 *Meal Recommendation*

Taking into account the need to produce, on a weekly basis, a meal plan adjusted to the needs of each facility, and mainly due to the fact that this is a very laborious process and whose responsibility would have to be entirely on the part of the social canteen, it would be necessary to produce a mechanism to automate this process.

Initially, each facility collects the needs of each of its patients, and makes a reservation with the social canteen, taking into account the nutritional needs of each patient, and also taking into account the age group of each of them. In addition, it is also necessary to take into account the type of food that each patient can have, which can be only liquid dishes, porridge, soups, among others...

Afterwards, the social canteen lists the requests made by the institution with the reservations made by the usual users of the social canteen, i.e., those who go to the canteen for lunch on a daily basis.

The number of meals varies greatly, and since it is necessary to carefully analyze the nutritional requirements of each meal, the process of creating the food plan for each week becomes a very time-consuming and repetitive task.

With this in mind, a module was created with the goal of automatically generating meal plans for each ward of the social center.

Despite the existence of the various modules mentioned above, the process of identifying the users, in order to understand the individual needs of each one is still very time consuming, since it is necessary to produce a food plan that meets the needs of each user individually, and for that it is necessary to analyse every week which users can eat certain foods, and which shouldn't.

With this in mind, a module was developed that still functions as a proof of concept, due to the inherent risks in generating food plans for such a large number of people.

This module uses an Machine Learning (ML) algorithm that is based on the nutritional requirements of each patient, as well as their age and palliative needs, thus generating a list of dishes that are most appropriate for this user.

Algorithm

Based on the needs presented at the beginning of the project, and also based on the needs that emerged as the project was being developed, the system would have to be able to obtain a specialized food plan, based on the needs of each user, and in addition, it would also have to be able to do so in a way adapted to the constant changes that would be subject, whether

they be the introduction of new foods, new dishes or even the entry or exit of some of the users in the system.

Apart from these requirements, it would be necessary to take into consideration the amount of data existing in the current system, since the dataset of users is reduced, and the amount of dishes inserted into the system is also reduced, so it would be imperative that the algorithm to be used was efficient for a low amount of data.

In an initial reading, the approach regarding the ML algorithm was studied. As the algorithm to be developed would be based on previously labelled data, namely data related to the existing relationship between diseases and food. In other words, the objective of the algorithm to be developed would be to collect the previously classified information and from this, classify new information with the intention of predicting the outcome to be obtained.

After the study carried out, it was possible to realize that due to the low quantity of existing information about this topic, it would be complex to use non-supervised algorithms, since it would be complicated to find sufficiently defined patterns in the obtained information, so that a decision could be taken.

Although it was necessary to exercise a process of classification of the previous data, this process was largely facilitated due to the existing data in the social canteen, which although not abundant, was well defined and divided. In addition, supervised algorithms tend to be more accurate, which is an important parameter in this situation, since we are dealing with patient feeding, which is an extremely delicate issue.

After having defined that a supervision algorithm would be used, it was necessary to define what type of task would be performed by the algorithm: Classification or Regression. As previously described, the aim of the system to be developed would be to obtain foods and dishes that could be appropriate for certain users, meaning that the problem consisted of assigning the dishes and foods to certain diseases or conditions, defining where each patient belongs and making the relationship between these two problems.

In this way, the algorithm would have to return a discrete value that represented the accuracy with which the classification was carried out, mapping between the user and the food that may be suitable for them. A regression algorithm would be used if it was necessary to predict a continuous value, which belongs to a domain not previously determined. In this way, the analysis of the results that would need to be obtained for the system to work determined that the value would have to be discrete. The algorithm would have to return something within the parameters of a ”you can eat this dish” or ”you cannot eat this dish” when introduced a user and a list of possible dishes.

The implemented model made use of the K Nearest Neighbour (KNN) algorithm. This is a simple ML algorithm that stores previously received data and tries to find a defined number of samples close to it. The main point of KNN is that it allows the number of samples to be set using a user-defined constant.

This algorithm is instance-based, meaning that no internal model development is initially done that would later be used for sample classification. On the contrary, instances of the initial data are stored. In this way, the training phase is non-existent, or practically non-existent. This also means that all the training data must necessarily be used in the thesis phase which makes this process more time consuming and heavier in terms of memory required.

The algorithm to be used, KNN, is executed based on the value chosen for the constant k. The search for the value of this constant is the main step in the creation of this algorithm and is what will define whether the algorithm is efficient and has a good relationship between the results obtained and the expenditure of time and memory.

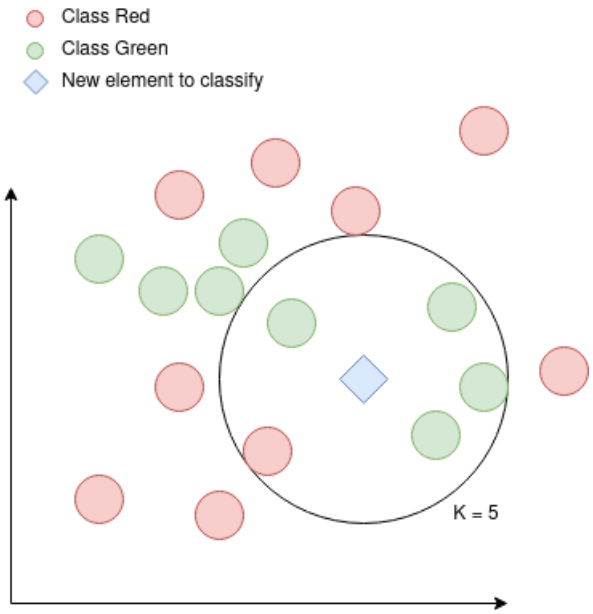


Figure 6: K-Nearest Neighbors example

When the number of classes in the system is 2, the value of k must be odd in order to obtain a value that is majority. Following the example of Figure 6, and with $k = 5$, the selection would fall on the green circles, since among the 5 selected instances, the majority is green.

Bearing in mind that this situation also occurs with a larger number of classes, a possibility that allows solving some of the problems that arise from it, and even allows obtaining values of k that might not allow finding a valid solution, would be to assign weights to the different neighbours, depending on the distance to which they are located. In this way, the closest instances will have a greater weight than the more distant instances.

Taking this into account, the value of k should be altered with the objective of obtaining the greatest balance between precision and computational cost, always bearing in mind that

an increase in precision always entails a greater computational cost, and a greater delay in finding the solution.

The decision to select the KNN has some advantages as well as disadvantages. The simplicity of the implementation is a point in favour of this choice, as well as the flexibility regarding the amount of features allowed.

Furthermore, the need to keep the algorithm updated and trained with the introduction of new classes is a fundamental point in favour of the KNN, since an algorithm that had to be pre-trained would cause problems in the long term, due to the need to re-train every time new data arises regarding which foods are recommended for certain health conditions, mainly due to the process in which this type of research is carried out, and which means that presently confirmed perceptions may not be the most appropriate in the near future.

The selection of this algorithm also brings adverse elements. The need to have to determine the value of the constant k is a time-consuming step, as described in the previous point, due to the need to balance computational cost with accuracy.

Even so, and regardless of the value of k, the computational cost of using this algorithm is always higher than other algorithms with pre-training, since this way, it is necessary to perform the calculation of the distance between the points in all queries performed, since this is, as described, an instance based algorithm that does not require a training phase.

The quality of the data is an important and decisive parameter in the quality of the results obtained when this algorithm is used. In this case, even though the amount of data is smaller, it is necessary to confirm that the data obtained has all the necessary parameters to create the representative axis. The speed is also one of the disadvantages of this algorithm when working with large amounts of data, which turns out not to be so unfavourable in this situation, because absurd amounts of information are not used in this case.

The developed algorithm goes through a data processing process, which generally consists of reading the data and respective treatment, and then the recommendation is made based on the processed data.

Processing

Data processing consists of reading a View of meals. This view allows obtaining a list of all the meals in a consistent way, taking into account that this query can remain the same regardless of the original tables from which the data is obtained. In other words, if in the future it is necessary to change any of the original tables, it won't be necessary to change the code, just update the view.

In the query for reading the meals, a list of the ingredients associated to each dish is also returned, together with the id of each ingredient. The reading of the necessary names is necessary to compare with the ingredients associated to the needs of each disease, while the ids are also necessary for the subsequent import of each associated nutrient.

In this way, after a correct reading of the meals in the database, a reading is made of the main nutrients in each of them.

After this process, it is necessary to understand which nutrients are recommended for each disease. It is essential to process the list of ingredients in each meal in order to standardise the ingredients obtained.

de — a — à — o — que — e — do — da — em — um — com

Figure 7: List of stop words (in Portuguese)

In other words, a dish with a list of ingredients with stop words or word connectors must be rewritten so that only the words referring to the ingredients are saved. Apart from the stop words, all other elements which may affect the reading of the sentence, such as parentheses, full stops, commas, etc., are also removed.

Before Normalization	**After Normalization**
Costeleta grelhada com arroz e batata frita	costeleta grelhada arroz batata frita
Filetes de Peixe (com arroz e batata frita)	filetes peixe arroz batata frita

Figure 8: Example of ingredients normalization

The processing of the collected data is fundamental for the data to be read from the database in a consistent and bullet-proof manner, so that any fault that may exist in the introduction of the data into the database is identified in the reading of the latter, and allows system failures to be avoided.

After the processing of the meals is completed, a dataframe is returned with a relationship between the foods and the diseases and/or health conditions for which they are recommended.

In addition to the data regarding the diseases, data obtained through the evaluations module, developed by a colleague for another dissertation, also in partnership with the social canteen, are also used, with which the users can carry out evaluations about the meals, thus allowing to have another point of analysis, which is important to provide more precise recommendations to the users. In this way, and since this system is not yet in use, dummy data were generated to the database, and through a view these are loaded in the processing phase, along with the data associated with the users.

Recommendation

The recommendation process consists of obtaining the profile of the user for whom the recommendation will be made. For this step, the user's profile must contain a series of

requirements that will be essential for the conclusion of this recommendation, such as the diet they are following, whether or not they are vegetarian or have some type of dietary restriction, as well as the diseases or health conditions they suffer from.

For the recommendation functions, the *sklearn* library is used, as it has a series of models and functions previously developed in order to facilitate the process of generation and use of **KNearest Neighbors**.

In addition to the profile of the user for whom the recommendation is to be obtained, it is also important to have a list of profiles for the other users. This list will be used to obtain data based on the other users of the system. Only in this way will it be possible to obtain valid results from the first moment the system is used.

This derives from the "cold start problem", where the small amount of parameterized data, in this case represented by the user's recent activity and their tastes, could mean that users will receive recommendations for dishes they do not like.

In this sense, the first step performed by the recommendation algorithm is to obtain a user-based recommendation, in order to exclude from the outset dishes that users with the same attributes as this one cannot eat, either due to dietary restrictions or health restrictions.

For the execution of this process, the features of the dataframe representative of the data regarding each user's profile are obtained.

The features are represented by the columns in this dataframe, and will be used later to position the users on the axis, as described in the explanation above regarding how the algorithm works.

With this in mind, the features to be obtained are the diseases/health conditions from which the user suffers and their dietary restriction. The dataframe obtained is similar to the example represented in the Figure 9.

calcium	carbohydrates	chloride	fiber	iron	vegan_diet	...
0	0	0	1	0	0	...
1	0	0	1	0	0	...
2	0	0	0	0	0	...
3	0	0	0	0	0	...
4	0	0	0	0	0	...

Figure 9: Example of Features Dataframe

Afterwards, it is necessary to find the neighbouring elements, which contain features similar to those of the user. This step will generate a dataframe that contains the values 1 or 0 associated to each column that each user has equal to the base user or not, respectively.

After the dataframe is created, the model with k neighbours is then created, and the "training phase" begins, that although this is not a training phase as it is usually done, it

is used to create data structures with the objective of maintaining the training dataset in a more organized way.

This process of searching for nearest neighbours can be a time consuming operation with larger datasets, and as such it may be necessary to optimise this phase. This is why the *fit()* function of the model is normally used so that the computational process does not have to be repeated for each instance, but can be performed once at this initial stage, and then reused. This function is also used to normalise the information and transform the data according to the *sklearn* API.

Subsequently, the nearest neighbours are obtained, returning a dataframe like the one represented in Figure 10.

id_user	Veg_Non	nutriente	diseases	dieta
8	non-veg	potássio	anemia obesidade colesterol	geral
1	non-veg	potássio	colesterol	high_fiber_diet geral
3	veg	sodium	obesidade	alkaline_diet
2	veg	protein	hipertensao	gluten_free_diet
5	veg	vitamin_a	anemia	high_protien_diet

Figure 10: Similar Users result dataframe

Once the process of obtaining similar users is concluded, it is necessary to understand which dishes they like.

At this point, the duplicate values are removed and the filtering of the dataframe obtained is carried out in order to keep only the columns necessary for the following steps.

These columns are the id of the dish, the name, the nutrient in greater quantity in this dish, the time of the year to which it has been associated and its ingredients.

The dish id is fundamental in this list, as it will be used later to remove duplicate dishes when joining the various dataframes obtained from the different types of recommendation. The nutrient in greater quantity can be used if you want to change the API in order to filter with more context, in other words, it may be important to obtain dishes with a higher percentage of a certain nutrient.

Regarding the time of year, this is one of the attributes associated to the dish when it is created, and it intends to separate the dishes by seasons of the year but not only, taking into account that, for example, the tendency is to have lighter meals in the summer and heavier ones in the winter. Even so, this parameter can always be adjusted according to the needs of the social canteen, and it is possible to add other values such as Christmas season, among others.

Finally, the value for ingredients returns the list of ingredients in its original form, before going through the normalisation process.

id_prato	nome	nutriente	epoca	ingredientes
6	costeleta grelhada com arro...	energia	inverno	arroz cozido, batata frita cas...
8	filetes de peixe com arro...	energia	verão	arroz cozido, batata frita cas...
9	paella	energia	verão	arroz com refogado, camarão...
10	arroz à valenciana	energia	verão	arroz à valenciana

Figure 11: Recommended meals result (after filter)

Once the dataframe corresponding to the recommended values based on the other users was obtained, it was necessary to obtain the data based on the recent activity of the user.

For this recommendation process, the module of evaluations described in the processing phase is used.

Initially, data regarding the recent activity of the user are obtained. From this recent activity, it is necessary to import, for each dish evaluated, the nutrient in greater quantity, the season, its ingredients, and the diet in which it is found.

Subsequently, a list is created with all the features obtained from these dataframes.

The next step, as in the process of obtaining neighbours based on other users, is to run the K-Nearest Neighbour algorithm with the developed dataframe.

With this algorithm it is again possible to find the dishes closest to the user on the Cartesian axis, that is, which dishes are most similar to the user's tastes, which in this case is the canteen user.

Once the recommended dishes have been obtained, the same filtering is performed as for the user-based recommendation, leaving the resulting dataset with columns representing the dish id, the name, the nutrient in greatest quantity in the dish, the time of year to which it is associated and its ingredients.

The next process consists of concatenating the dataframes obtained, removing duplicate values and resetting the indicators in this list.

id_prato	nome	nutriente	epoca	ingredientes
6	costeleta grelhada com...	energia	inverno	arroz cozido simples, batata fr...
8	filetes de peixe com...	energia	verão	arroz cozido simples, batata fr...
9	paella	energia	verão	arroz com refogado, camarã...
10	arroz à valenciana	energia	verão	arroz à valenciana
11	caldo verde	sal	inverno	sopa, caldo verde
12	tarte de maçã	energia	verão	tarte de maçã
5	costeleta grelhada com arroz	energia	inverno	arroz cozido simples, borr...
13	ameijoas à bulhão pato	energia	inverno	amêijoas à bulhão pato

Figure 12: Recommended meals result (after filter)

This way, and having all the unique dishes obtained, it is possible to obtain their ids, being able to advance to the next phase, which consists of distributing the obtained meals by the several days of the week.

7.4.5 *Weekly Meal Plan Generation*

To enable the weekly organisation of menus, a module was developed using Linear Programming (LP) that distributes the recommended meals throughout the week, according to the caloric needs and dietary requirements of the users, such as the obligation to have breakfast, lunch and dinner every day.

In this way, the recommended meals for each user are generated using the previously mentioned module (Meal Recommendation), followed by their distribution for each day of the week. The current module can work independently, although it allows for greater efficiency when used in conjunction with the recommendation module.

In this step, it is essential to obtain the necessary data, not only regarding the meals but also the foods that compose them, since the objective of this process is to distribute the protein and calorie cost of each meal throughout the week, that is, it is mandatory to have all the data available regarding meals and foods.

Subsequently, the functions in which the information is processed were defined, in which case the caloric and protein needs of the user for whom the meals are to be obtained are obtained. Besides these needs, it is also possible to change the number of meals that the user must consume daily, since this value can vary according to the diet and needs of each person.

The meals are separated according to their characteristics, which may correspond to soup, breakfast, dessert or main course (besides the already defined ”main courses”, salads will also be inserted in this category). This separation process makes it possible to create

complete meals, and to avoid that needs are only met through unusual variations, such as the consumption of multiple main courses in the same meal. Only using this separation by category allows the creation of the restrictions that make the meal distribution process possible.

Constraints

In the context of linear programming, it was defined that the objective of the function is to maximise the number of calories, taking into account that it is always intended that the calorie needs of the user are achieved daily.

In relation to the restrictions, it is intended that the number of daily calories be between the values defined for the selected user and it is intended to have an exact number of soups, main dishes and desserts each day, and this number can be defined for each user.

Algorithm

Solving the problem that aims to obtain the food distribution is one of the most time consuming stages of this recommendation process, and as such two algorithms were developed that have the same basis and aim to return similar results but use different Python libraries. The first algorithm developed uses *pywraplp* and the second algorithm developed uses the *pulp* library.

Both algorithms start from the same basic principle of linear programming. Initially the objective function is established and what is to be achieved. Afterwards the constraints are registered, and finally the solver is run, which executes the function with the objective of understanding which values can be used for the variables that you want to achieve.

In this situation, the solver was executed for each day, and the food restrictions are defined for each day, that is, the algorithm is executed for the 7 days of each week, and the meals obtained each day are registered, in order to avoid repeating meals in the same week. Besides, a randomness factor is assigned to each day, in order to avoid that the result is the same every week, and taking into account that the algorithm aims to find the optimal value of the problem, this would always be the same if all constraints and constants were the same.

In the case of the algorithm developed using pywraplp, on each day the constraints are defined using an object with the format $\{\text{'Prato Principal': 2, 'Sopa': 2, 'Sobremesa': 1}\}$. Based on this constraint, X possibilities of meals are generated. The Solver is then created, along with the necessary constraints for the user.

With this in mind, a value representing the quality of the data obtained is generated, meaning how far the calories and protein obtained in the combination of dishes generated are from the objectives that are intended to be achieved. This value is associated with each solution. In this algorithm, the linear programming problem is used with the intent of

minimizing the obtained value, taking into account that this value is the existing error in the presented solution.

This list that contains all the solutions is then sorted with the intuition of obtaining only the solutions with the smallest error, removing the excess and keeping the most adequate results for that particular day.

A list is developed with the weights of each solution, where the solutions with the highest weights have more possibilities of being used. This way it is possible to increase the randomness of the algorithm, but still maintaining its accuracy. The function choices() from the random library allows a random result to be obtained when the list of solutions and the weight of each of these solutions is presented.

[0.019801980198019802, 0.019603960396039604, 0.019405940594059406,
0.019207920792079208, 0.01900990099009901, 0.01881188118811881,
0.018613861386138613, 0.018415841584158415, 0.018217821782178217,
0.01801980198019802, 0.01782178217821782, 0.017623762376237622,
0.017425742574257427, 0.01722772277227723, ...]

This implementation with *pywraplp* is executed once for each day of the week. Still, it is possible to reduce the execution time of this implementation by making use of the function choices() to distribute the solutions by the days of the week, since the function allows passing a value *k* that represents the number of results that is intended to be obtained.

Even so, this possibility became unfeasible due to the need to remove the previously selected dishes, which using the choices() function would not be possible.

In the case of the algorithm developed using *pulp*, on each day the necessary constraints are defined, as in the implementation of *pywraplp*.

The implementation of the two algorithms is similar, but the difference lies essentially in the objective selected when using *purp*. While in the previous implementation an error margin that each solution presents in relation to the intended objectives is calculated, and the objective of the function is to find the solution with the minimum value for that variable, in this case the objective of the function is to find the solution in which the sum of calories of all the meals is maximum, within the imposed restrictions.

In other words, this guarantees that the solution corresponds to the user's needs in terms of the number of calories required.

Using *purp*, a loop is executed for each day of the week, and the dishes are inserted in the correct category (Main dish, soup, dessert, breakfast), and if a dish has already been selected for a day of the same week, it is not inserted in any category, being ignored.

The solver is then created with the objective function, which in this case is the maximization of the number of calories, and the respective restrictions. Below you can see an example of the restrictions that are inserted in the solver of *purp*.

_**C1**: $85 * idMeal1 + 11 * 1idMeal0 + 89 * idMeal3 + 27 * idMeal5 + 50 * idMeal6 + 65 * idMeal8 + 442 * idMeal9 >= minCalories$

_**C2**: $85 * idMeal1 + 11 * 1idMeal0 + 89 * idMeal3 + 27 * idMeal5 + 50 * idMeal6 + 65 * idMeal8 + 442 * idMeal9 <= maxCalories$

_**C3**: $count(idMealDessert) = 1$

_**C4**: $count(idMealX) + count(idMealX) + count(idMealX) + count(idMealX) + count(idMealX) + count(idMealX) <= 2$

_**C5**: $count(idMealSoup) = 2$

_**C6**: $count(idMealBreakfast) = 1$

The _**C1** constraint indicates that the sum of the calories of the selected meals must be greater than the value used for the minimum calories, while the _**C2** constraint indicates that this same sum must be less than the maximum calories defined.

On the other hand, the _**C3**, _**C5** and _**C6** restrictions are restrictions that are used to ensure that dessert, soup, and breakfast are used exactly once, twice and once respectively, while the _**C4** restriction represents that the number of main meals in the same day must not exceed two.

The big difference in the implementation using the *pywraplp*, is that if there is a solution where the sum of the calories of the meals is at the lower edge of the constraint, or the sum of the calories is extremely low but just above the minimum limit, and there is another more suitable solution with a higher sum, but that sum is further away from the maximum limit (the preferential limit), then the selected solution would be the incorrect one, which does not happen in this implementation with the *purp*.

In addition to this difference, the execution of the algorithm with the *purp* is faster, because it does not require that the weight of all solutions be calculated in its execution.

The results were developed and normalised so that the requests to the API are returned with the same structure, this way it is possible to test the results obtained with the different algorithms and the different implementations of them.

7.5 TEST CASE DEVELOPMENT

As a way to demonstrate the development of this recommendation solution, the previously developed dummy data that is used as the basis for this current system will be used.

It is fundamental to take into account that this data will be later removed and replaced by real data (when available), and that if there are no real cases, only data directly related to meals will be used.

In addition to this, and this being a proof of concept, it is also important to take into consideration that this system should not be used in real environments, mainly due to the

lack of tests carried out regarding the links between dishes and diseases, as this is a process that may require a much larger amount of data if sufficiently accurate results are to be obtained in a high risk environment, such as that of meal generation for users with enhanced healthcare.

Once the example user, represented in Figure 13, has been chosen, it is possible to understand the diseases this user may have, the dietary restrictions and their diet.

idUser	food restrictions	diseases	diet
8	geral	anemia obesidade colesterol	geral

Figure 13: Example of Features Dataframe

With this base, and making the request to the recommendation API for this user's recommendations, the JSON object with the following structure is returned:

```
{
    "meals": [
        { "prato": 11, "quantidade": 2 },
        { "prato": 6, "quantidade": 1 },
        { "prato": 8, "quantidade": 1 },
        { "prato": 15, "quantidade": 3 }
    ],
    (...)
}
```

This list of meals obtained represents Caldo Verde (11), Costeleta grelhada com com arroz e batata frita (6), Filetes de peixe com arroz e batata frita (8) e Papas de Aveia (15).

So that the results could be viewed directly through the web application, a recommendations page was introduced, through which managers can view the recommendations obtained for a particular week for the user they select.

7.6 Results Discussion

The results obtained confirm what was already expected, namely that it is necessary to take into account that the recommendation module of this system is for now only a proof of concept, bearing in mind that the margin of error is still too high for what is one of the bases of health, which is food, especially in patients who may be in a fragile state of health and may need extra health care.

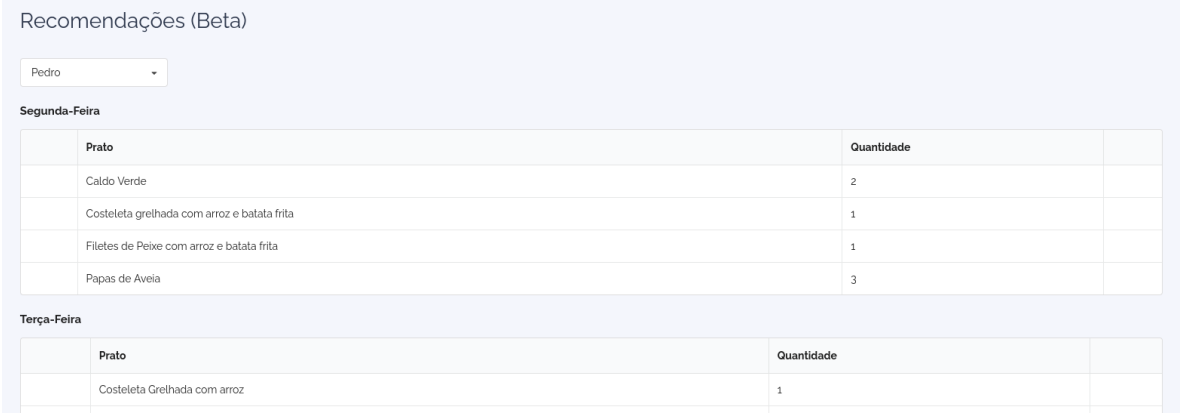


Figure 14: User Recommendation Example on Web App

Despite the results obtained indicating that the calorie and protein level restrictions are met, it would also be necessary to take into account the remaining macro- and micro-nutritional values, considering the need for a balanced diet, which means that it would also be necessary to take into account the remaining nutrients.

In addition, the categorization and distribution of dishes throughout the day may not be the most appropriate according to the results obtained, since some of the restrictions were not met in some of the results, such as the need to have exactly two main dishes, which did not happen in some of the situations in which the algorithm associated only one or even more than two main dishes for the same day.

Even so, it is also important to take into account that the distribution of meals over several weeks may be a process made difficult by the small number of dishes in the system, and by the fact that some of the dishes inserted are not used by the social canteen and as such may not correspond correctly to the existing one. Thus, the components in each of these dishes may not correspond to those used and so the total value of the nutrients may not be as expected. In other words, the existence of incorrect components, or lack of them, in each of the dishes created may be the basis of the results obtained, since the final calculations may be compromised by the total calories or proteins in each dish.

Still, the results are demonstrative of a solution with a future, which needs some improvements but that with the increase of the existing data could be taken into consideration as an adequate solution for the indicated requirements, or even could be currently implemented to serve as a basis or inspiration in the choice of the weekly meal plan for each of the wards.

8

MAIN DIFFICULTIES

8.1 PROJECT MANAGEMENT

The necessary requirements for the project were not defined from the beginning of the development, and it was a long process, due to the difficulty of communication with the entity for which the system was intended to be developed.

In this way, the initial development of the project ended up being delayed and with that all the steps and objectives initially defined ended up not being fulfilled in the foreseen dates.

This situation allied to the existence of a pandemic, which required that all the work was done from home made it even more difficult to understand what would be necessary for this system, especially taking into account that the communication was made with those responsible for a hospital/health centre, and as such would have reduced availability considering the circumstances encountered.

Furthermore, the requirements presented were always dubious, and some of the needs initially presented as essential ended up being denied, especially considering that the data necessary for the development of some of these functionalities were not made available by the entity, as is the case of the integration between the web page and the invoicing program.

8.2 AVAILABLE DATA

The main problem in developing this solution was to obtain structured and updated data that could be used for the correct functioning of the system. The search for this information always became complicated and time-consuming, and the final decision fell on the development of a dataset, created by joining information obtained from various sources, and that would be used in the development of the system.

The production of this dataset was an extremely time consuming process, taking more time than expected to develop it, so it was necessary to try to obtain this information in a

more automatic and less laborious way, but it was only in this way that it was possible to develop the desired algorithm.

Therefore, this list of data was obtained in a somewhat rudimentary and manual way, thus making the process of maintaining and updating this data quite complex.

In addition to the data used for the development of the algorithm, it was also complicated to obtain the data relating to the recipes used by the social canteen, and, as described, it was necessary to generate some recipes from the existing ones so that the system could be tested outside the production environment. This was mainly due to the way the data was stored in the system in place in the social canteen, where it was quite complicated to access it.

9

CONCLUSION AND FUTURE WORK

With the evolution of technology, its use linked to health systems is in an upward phase. When applied in this area, information technology systems can improve the well-being and health of the users of these establishments. In addition, machine learning algorithms have also been increasingly used in this type of systems, namely in the production of recommendation systems that can be used in decision support.

In this way, and considering what was described above, this dissertation project arose from a partnership between Universidade do Minho and the Vila Verde Social Canteen.

This project starts from the need to automate the process of generating meals in the canteen and the social kitchen. After an analysis and study of the existing projects and algorithms, it was possible to outline which strategy to follow in order to produce a platform that allows planning the meals of the users based on their nutritional needs and preferences.

The next step of this dissertation is to develop this system and implement it in the social canteen, applying the knowledge obtained throughout the project.

The future work that could be implemented in this solution lies in a series of procedures that can be used to improve the way data is obtained and used, and also with the aim of improving the viability of the system so that it can be used in practice as a way of generating the meals of the social canteen.

9.1 AUTOMATE PROCESS OF OBTAINING DATA

As described in the Chapter of the main difficulties, the process of obtaining data on diseases and food is quite time-consuming and laborious, since it is necessary to continue researching new studies on what is recommended in this association, and as such it is a process that should be improved in order to make this solution practical.

One of the solutions that could be used to solve this problem would be the use of Natural Language Processing (NPL) in order to read the new published studies and understand what data can be obtained that makes sense in this context. Even so, this solution may also be unfeasible due to the need to understand which studies are marked as valid, or which previously obtained information is already defined as invalid due to new studies done.

9.2 IMPLEMENT INVOICE INTEGRATION

As this was one of the initial requirements presented, and which would end up not being developed due to the fact that the necessary data to implement this module had not been provided, it was essential that one of the steps to be followed in the development of this solution was to carry out an integration with the existing invoicing system of the social canteen.

Along with the mobile application developed for the same entity in another project, this integration system could be used not only to generate the users' invoices when they pay their meals, but it could also be used to read the existing stock of certain foods, and thus improve the algorithm in order to generate meals only with existing ingredients, if necessary.

9.3 GENERATE MEAL PLAN FOR WARD

Currently the generation of a weekly meal plan is done directly for each particular user. Since this is impractical and it is not possible to generate a menu per person, one of the fundamental steps to be taken would be to modify the algorithm so that it can generate a global menu based on the reservations that have been made.

For this to happen, it would also be necessary to modify the reservations module so that each person in charge of the services could indicate which users would be present in the reservations made, and thus obtain a list of those meals that would be the most recommended for as many users as possible.

10

ACKNOWLEDGEMENTS

This thesis represents the culmination of a season of hard work and study aimed at developing a system that would meet the expectations placed at the beginning of this campaign. This goal would not have been achievable without the support of those who were present throughout the process, and who contributed to making it possible.

First of all, I would like to express my gratitude to my supervisor Professor José Machado, for giving me this opportunity, for presenting me with this challenging project and for providing the necessary support throughout the last few months.

I would also like to thank Professor António Abelha for the support provided throughout the development of the thesis, and for having, along with Professor José Machado, given me the opportunity to participate in this project.

I would also like to extend my appreciation to Santa Casa da Misericórdia de Vila Verde for their availability in receiving us for this project, and also for their cooperation and for providing everything that we needed and that was within their reach.

This project would not have been possible without Centro ALGORITMI, especially Rui Miranda, to whom I am truthfully grateful for his availability to help me whenever necessary, for his patience and dedication, and for the support provided throughout this year, especially considering the circumstances in which it took place.

Relatively to my academic formation, a word of gratitude to Instituto Politécnico de Viana do Castelo (IPVC) that received me during my graduation in Computer Engineering and to Universidade do Minho for having received me for my Masters.

Finally, the biggest appreciation goes to my family and friends, especially my parents, my brother, and my girlfriend, for giving me the necessary strength to continue to face the adversities that appeared throughout these years.

BIBLIOGRAPHY

[1] Joyce Green Pastors, Hope Warshaw, Anne Daly, Marion Franz, and Karmeen Kulkarni. The evidence for the effectiveness of medical nutrition therapy in diabetes management. *Diabetes care*, 25:608–613, 2002. ISSN 01495992. doi: 10.2337/diacare.25.3.608.

[2] Nadja Leipold, Mira Madenach, Hanna Schäfer, Martin Lurz, Nada Terzimehic, Georg Groh, Markus Böhm, Kurt Gedrich, and Helmut Krcmar. Nutrilize a personalized nutrition recommender system: An enable study. *CEUR Workshop Proceedings*, 2216: 24–29, 2018. ISSN 16130073.

[3] Bo Ram Wang, Ji Yun Park, Kyungyong Chung, and In Young Choi. Influential factors of smart health users according to usage experience and intention to use. *Wireless Personal Communications*, 79:2671–2683, 2014. ISSN 09296212. doi: 10.1007/s11277-014-1769-0.

[4] Nikos Manouselis, Katrien Verbert, Hendrik Drachsler, and Olga C. Santos. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17:377, 2005. doi: 10.1145/1864708.1864797.

[5] José Neves, Henrique Vicente, Marisa Esteves, Filipa Ferraz, António Abelha, José Machado, Joana Machado, João Neves, Jorge Ribeiro, and Lúzia Sampaio. A deep-big data approach to health care in the ai age. *Mobile Networks and Applications*, 23: 1123–1128, 2018. ISSN 15728153. doi: 10.1007/s11036-018-1071-6.

[6] Marc Wieland and Massimiliano Pittore. Performance evaluation of machine learning algorithms for urban pattern recognition from multi-spectral satellite images. *Remote Sensing*, 6:2912–2939, 2014. ISSN 20724292. doi: 10.3390/rs6042912.

[7] Li-Pang Chen. Mehryar mohri, afshin rostamizadeh, and ameet talwalkar: Foundations of machine learning, second edition. *Statistical Papers*, 60:1793–1795, 2019. ISSN 0932-5026. doi: 10.1007/s00362-019-01124-9. URL https://doi.org/10.1007/s00362-019-01124-9.

[8] Ivens Portugal, Paulo Alencar, and Donald Cowan. The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications*, 97:205–227, 2018. ISSN 09574174. doi: 10.1016/j.eswa.2017.12.020. URL https://doi.org/10.1016/j.eswa.2017.12.020.

[9] S. Sumathi and S. N. Sivanandam. Introduction to data mining principles, 2007. ISSN 1860949X.

[10] Paulo Cortez and Alice Silva. Using data mining to predict secondary school student performance. *15th European Concurrent Engineering Conference 2008, ECEC 2008 - 5th Future Business Technology Conference, FUBUTEC 2008*, 2003:5–12, 2008.

[11] Behrouz Minaei-Bidgoli, Deborah A. Kashy, Gerd Kortemeyer, and William F. Punch. Predicting student performance: An application of data mining methods with an educational web-based system. *Proceedings - Frontiers in Education Conference, FIE*, 1: T2A13–T2A18, 2003. ISSN 15394565. doi: 10.1109/FIE.2003.1263284.

[12] Pedro Domingos. A few useful things to know about machine learning. *communications of the ACM*, 55:79–88, 2012. URL https://dl.acm.org/citation.cfm?id=2347755.

[13] Machine learning: What it is and why it matters.

[14] Fabio Fabris, João Pedro de Magalhães, and Alex A. Freitas. A review of supervised machine learning applied to ageing research. *Biogerontology*, 18:171–188, 2017. ISSN 15736768. doi: 10.1007/s10522-017-9683-y.

[15] Taiwo Oladipupo. Types of machine learning algorithms. *New Advances in Machine Learning*, 2010. doi: 10.5772/9385.

[16] Muhammad Usama, Junaid Qadir, Aunn Raza, Hunain Arif, Kok Lim Alvin Yau, Yehia Elkhatib, Amir Hussain, and Ala Al-Fuqaha. Unsupervised machine learning for networking: Techniques, applications and research challenges. *IEEE Access*, 7: 65579–65615, 2019. ISSN 21693536. doi: 10.1109/ACCESS.2019.2916648.

[17] Machine learning, deep learning 101. URL https://developer.ibm.com/articles/l-machine-learning-deep-learning-trs/.

[18] Alex Gammerman, Volodya Vovk, and Vladimir Vapnik. Learning by transduction. 2013. URL http://arxiv.org/abs/1301.7375.

[19] G Kesavaraj and S Sukumaran. A study on classification techniques in data mining. 2013.

[20] Newsom Usp and S E M Winter. Latent variables, 2020.

[21] Pratiyush Guleria and Manu Sood. *Intelligent Learning Analytics in Healthcare Sector Using Machine Learning*. 2020. ISBN 9783030408497. doi: 10.1007/978-3-030-40850-3_3.

[22] Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B Kantor, and Francesco Ricci. *Recommender Systems Handbook*. 2011. ISBN 9780387858197. doi: 10.1007/978-0-387-85820-3.

[23] Hoill Jung and Kyungyong Chung. Knowledge-based dietary nutrition recommendation for obese management. *Information Technology and Management*, 17:29–42, 2016. ISSN 15737667. doi: 10.1007/s10799-015-0218-4.

[24] The role of mathematical models in operations decision making b2 constrained optimization models b2 advantages and disadvantages of using optimiza-tion models b5 assumptions of linear programming models b6 formulating linear programs b7 the geometry of linear programs b14 the graphical solution approach b15 the simplex algorithm b17 using artificial variables b26 computer solutions of linear programs b29 using linear programming models for decision making b32.

[25] Michel Goemans. Linear programming, 2015.

[26] Giuseppe Agapito, Mariadelina Simeoni, Barbara Calabrese, Ilaria Caré, Theodora Lamprinoudi, Pietro H. Guzzi, Arturo Pujia, Giorgio Fuiano, and Mario Cannataro. Dietos: A dietary recommender system for chronic diseases monitoring and management. *Computer Methods and Programs in Biomedicine*, 153:93–104, 2018. ISSN 18727565. doi: 10.1016/j.cmpb.2017.10.014.

[27] Rodrigo Zenun Franco. Online recommender system for personalized nutrition advice. *RecSys 2017 - Proceedings of the 11th ACM Conference on Recommender Systems*, pages 411–415, 2017. doi: 10.1145/3109859.3109862.

[28] Vanesa Espín, María V. Hurtado, and Manuel Noguera. Nutrition for elder care: A nutritional semantic recommender system for the elderly. *Expert Systems*, 33:201–210, 2016. ISSN 14680394. doi: 10.1111/exsy.12143.

[29] Longqi Yang, Cheng Kang Hsieh, Hongjian Yang, John P. Pollak, Nicola Dell, Serge Belongie, Curtis Cole, and Deborah Estrin. Yum-me: A personalized nutrient-based meal recommender system. *ACM Transactions on Information Systems*, 36, 2017. ISSN 15582868. doi: 10.1145/3072614.

[30] Bill Kuechler and Stacie Petter. D esign s cience r esearch in. pages 1–66, 2012. doi: 1756-0500-5-79[pii]\r10.1186/1756-0500-5-79. URL http://www.desrist.org/design-research-in-information-systems/[Accessed11may2017].

[31] Ken Peffers, Tuure Tuunanen, Charles E. Gengler, Matti Rossi, Wendy Hui, Ville Virtanen, and Johanna Bragge. The design science research process: A model for producing and presenting information systems research. *arXiv*, 2006. ISSN 23318422.

[32] Nimesh Chhetri. A comparative analysis of node . js (server-side javascript). *Culminating Projects in Computer Science and Information Technology*, page 79, 2016.

[33] Express.js. Express - node.js web application framework, 2017. URL https://expressjs.com/.

[34] White Paper. Creating rest apis to enable our connected world. 2016.

[35] React. React a javascript library for building user interfaces, 2016. URL https://reactjs.org/.

[36] Mayelson De Sousa. humanportal – a react . js case study. pages 24–27, 2020.

[37] Dan Abramov and Redux documentation authors. Redux - a predictable state container for javascript apps., 2020. URL https://redux.js.org/.

[38] MDN. Django introduction - mdn. URL https://developer.mozilla.org/en-US/docs/Learn/Server-side/Django/Introduction.

[39] Django Foundation. Django project, 10 2021. URL https://www.djangoproject.com/.

[40] Yoji Kato, Tokio Domoto, Masanori Hiramitsu, Takao Katagiri, Kimiko Sato, Yukiko Miyake, Satomi Aoi, Katsuhide Ishihara, Hiromi Ikeda, Namiko Umei, Atsusi Takigawa, and Toshihide Harada. Effect on blood pressure of daily lemon ingestion and walking. *Journal of Nutrition and Metabolism*, 2014, 2014. ISSN 20900732. doi: 10.1155/2014/912684. URL /pmc/articles/PMC4003767//pmc/articles/PMC4003767/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4003767/.

[41] Sedigheh Asgary and Mahtab Keshvari. Effects of citrus sinensis juice on blood pressure. *ARYA Atherosclerosis*, 9:98, 2013. ISSN 22516638. URL /pmc/articles/PMC3653258//pmc/articles/PMC3653258/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3653258/.

[42] Mark G. Filipovic, Stefanie Aeschbacher, Martin F. Reiner, Simona Stivala, Sara Gobbato, Nicole Bonetti, Martin Risch, Lorenz Risch, Giovanni G. Camici, Thomas F. Luescher, Clemens Von Schacky, David Conen, and Juerg H. Beer. Whole blood omega-3 fatty acid concentrations are inversely associated with blood pressure in young, healthy adults. *Journal of Hypertension*, 36:1548, 7 2018. ISSN 14735598. doi: 10.1097/HJH.0000000000001728. URL /pmc/articles/PMC6085127//pmc/articles/PMC6085127/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6085127/.

[43] P. Xun, N. Hou, M. Daviglus, K. Liu, J. S. Morris, J. M. Shikany, S. Sidney, D. R. Jacobs, and K. He. Fish oil, selenium and mercury in relation to incidence of hypertension: a 20-year follow-up study. *Journal of internal medicine*, 270:175–186, 8 2011. ISSN 1365-2796. doi: 10.1111/J.1365-2796.2010.02338.X. URL https://pubmed.ncbi.nlm.nih.gov/21205024/.

[44] Mark C. Houston. The importance of potassium in managing hypertension. *Current hypertension reports*, 13:309–317, 8 2011. ISSN 1534-3111. doi: 10.1007/S11906-011-0197-8. URL https://pubmed.ncbi.nlm.nih.gov/21403995/.

[45] Mark Houston. The role of magnesium in hypertension and cardiovascular disease. *Journal of Clinical Hypertension*, 13:843–847, 11 2011. ISSN 15246175. doi: 10.1111/J.1751-7176.2011.00538.X.

[46] Marc P. McRae. Therapeutic benefits of l-arginine: An umbrella review of meta-analyses. *Journal of Chiropractic Medicine*, 15:184, 2016. ISSN 15563707. doi: 10.1016/J.JCM.2016.06.002. URL /pmc/articles/PMC5021928//pmc/articles/PMC5021928/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5021928/.

[47] Vincenzo Savica, Guido Bellinghieri, and Joel D. Kopple. The effect of nutrition on blood pressure. *Annual review of nutrition*, 30:365–401, 8 2010. ISSN 1545-4312. doi: 10.1146/ANNUREV-NUTR-010510-103954. URL https://pubmed.ncbi.nlm.nih.gov/20645853/.

[48] Alexei Wong, Danielle Viola, Douglas Bergen, Eileen Caulfield, J. Mehrabani, and Arturo Figueroa. The effects of pumpkin seed oil supplementation on arterial hemodynamics, stiffness and cardiac autonomic function in postmenopausal women. *Complementary therapies in clinical practice*, 37:23–26, 11 2019. ISSN 1873-6947. doi: 10.1016/J.CTCP.2019.08.003. URL https://pubmed.ncbi.nlm.nih.gov/31445363/.

[49] Viranda H. Jayalath, Russell J. De Souza, John L. Sievenpiper, Vanessa Ha, Laura Chiavaroli, Arash Mirrahimi, Marco Di Buono, Adam M. Bernstein, Lawrence A. Leiter, Penny M. Kris-Etherton, Vladimir Vuksan, Joseph Beyene, Cyril W.C. Kendall, and David J.A. Jenkins. Effect of dietary pulses on blood pressure: A systematic review and meta-analysis of controlled feeding trials. *American Journal of Hypertension*, 27:56, 1 2014. ISSN 08957061. doi: 10.1093/AJH/HPT155. URL /pmc/articles/PMC5391775//pmc/articles/PMC5391775/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5391775/.

[50] Stefano Vendrame and Dorothy Klimis-Zacas. Potential factors influencing the effects of anthocyanins on blood pressure regulation in humans: A review. *Nutrients*, 11, 6 2019. ISSN 20726643. doi: 10.3390/NU11061431. URL /pmc/articles/PMC6628116//pmc/articles/PMC6628116/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6628116/.

[51] Lukas Schwingshackl, Carolina Schwedhelm, Georg Hoffmann, Sven Knüppel, Khalid Iqbal, Violetta Andriolo, Angela Bechthold, Sabrina Schlesinger, and

Heiner Boeing. Food groups and risk of hypertension: A systematic review and dose-response meta-analysis of prospective studies. *Advances in Nutrition*, 8:793, 11 2017. ISSN 21565376. doi: 10.3945/AN.117.017178. URL /pmc/articles/PMC5683007//pmc/articles/PMC5683007/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5683007/.

[52] Katherine A. Sauder, Cindy E. McCrea, Jan S. Ulbrecht, Penny M. Kris-Etherton, and Sheila G. West. Pistachio nut consumption modifies systemic hemodynamics, increases heart rate variability, and reduces ambulatory blood pressure in well-controlled type 2 diabetes: a randomized trial. *Journal of the American Heart Association: Cardiovascular and Cerebrovascular Disease*, 3, 2014. ISSN 20479980. doi: 10.1161/JAHA.114.000873. URL /pmc/articles/PMC4310367//pmc/articles/PMC4310367/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4310367/.

[53] Noushin Mohammadifard, Amin Salehi-Abargouei, Jordi Salas-Salvadó, Marta Guasch-Ferré, Karin Humphries, and Nizal Sarrafzadegan. The effect of tree nut, peanut, and soy nut consumption on blood pressure: a systematic review and meta-analysis of randomized controlled clinical trials. *The American Journal of Clinical Nutrition*, 101:966–982, 5 2015. ISSN 0002-9165. doi: 10.3945/AJCN.114.091595. URL https://academic.oup.com/ajcn/article/101/5/966/4577579.

[54] Andrew S. Potter, Shahrzad Foroudi, Alexis Stamatikos, Bhimanagouda S. Patil, and Farzad Deyhim. Drinking carrot juice increases total antioxidant status and decreases lipid peroxidation in adults. *Nutrition Journal*, 10:96, 2011. ISSN 14752891. doi: 10.1186/1475-2891-10-96. URL /pmc/articles/PMC3192732//pmc/articles/PMC3192732/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3192732/.

[55] Narges Hedayati, Mehri Bemani Naeini, Arash Mohammadinejad, and Seyed Ahmad Mohajeri. Beneficial effects of celery (apium graveolens) on metabolic syndrome: A review of the existing evidences. *Phytotherapy research : PTR*, 33:3040–3053, 12 2019. ISSN 1099-1573. doi: 10.1002/PTR.6492. URL https://pubmed.ncbi.nlm.nih.gov/31464016/.

[56] Q. Chan, J. Stamler, I. J. Brown, M. L. Daviglus, L. Van, A. R. Dyer, L. M. Oude Griep, K. Miura, H. Ueshima, L. Zhao, J. K. Nicholson, E. Holmes, and P. Elliott. Relation of raw and cooked vegetable consumption to blood pressure: the intermap study. *Journal of Human Hypertension*, 28:353, 2014. ISSN 14765527. doi: 10.1038/JHH.2013.115. URL /pmc/articles/PMC4013197//pmc/articles/PMC4013197/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4013197/.

[57] Ioana Mozos, Dana Stoian, Alexandru Caraba, Clemens Malainer, Jaroslaw O. Horbanczuk, and Atanas G. Atanasov. Lycopene and vascular health. *Frontiers in*

Pharmacology, 9:521, 5 2018. ISSN 16639812. doi: 10.3389/FPHAR.2018.00521. URL /pmc/articles/PMC5974099//pmc/articles/PMC5974099/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5974099/.

[58] Ho Ming Cheng, Georgios Koutsidis, John K. Lodge, Ammar Ashor, Mario Siervo, and José Lara. Tomato and lycopene supplementation and cardiovascular risk factors: A systematic review and meta-analysis. *Atherosclerosis*, 257:100–108, 2 2017. ISSN 0021-9150. doi: 10.1016/J.ATHEROSCLEROSIS.2017.01.009. URL http://www.atherosclerosis-journal.com/article/S0021915017300102/fulltexthttp://www.atherosclerosis-journal.com/article/S0021915017300102/abstracthttps://www.atherosclerosis-journal.com/article/S0021-9150(17)30010-2/abstract.

[59] Jaime L. Clark, Peter Zahradka, and Carla G. Taylor. Efficacy of flavonoids in the management of high blood pressure. *Nutrition reviews*, 73:799–822, 12 2015. ISSN 1753-4887. doi: 10.1093/NUTRIT/NUV048. URL https://pubmed.ncbi.nlm.nih.gov/26491142/.

[60] Lea Borgi, Isao Muraki, Ambika Satija, Walter C. Willett, Eric B. Rimm, and John P. Forman. Fruit and vegetable consumption and the incidence of hypertension in three prospective cohort studies. *Hypertension (Dallas, Tex. : 1979)*, 67:288, 2 2016. ISSN 15244563. doi: 10.1161/HYPERTENSIONAHA.115.06497. URL /pmc/articles/PMC5350612//pmc/articles/PMC5350612/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5350612/.

[61] Haruki Nakamura, Hiromasa Tsujiguchi, Akinori Hara, Yasuhiro Kambayashi, Sakae Miyagi, Thao Thi Thu Nguyen, Keita Suzuki, Yuichi Tao, Yuriko Sakamoto, Yukari Shimizu, Norio Yamamoto, and Hiroyuki Nakamura. Dietary calcium intake and hypertension: Importance of serum concentrations of 25-hydroxyvitamin d. *Nutrients*, 11, 4 2019. ISSN 20726643. doi: 10.3390/NU11040911. URL /pmc/articles/PMC6521038//pmc/articles/PMC6521038/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6521038/.

[62] Lukas Schwingshackl, Carolina Schwedhelm, Georg Hoffmann, Sven Knüppel, Khalid Iqbal, Violetta Andriolo, Angela Bechthold, Sabrina Schlesinger, and Heiner Boeing. Food groups and risk of hypertension: A systematic review and dose-response meta-analysis of prospective studies. *Advances in Nutrition*, 8:793, 11 2017. ISSN 21565376. doi: 10.3945/AN.117.017178. URL /pmc/articles/PMC5683007//pmc/articles/PMC5683007/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5683007/.

[63] Kate S. Driscoll, Amanda Appathurai, Markandeya Jois, and Jessica E. Radcliffe. Effects of herbs and spices on blood pressure: a systematic literature review of randomised

controlled trials. *Journal of hypertension*, 37:671–679, 4 2019. ISSN 1473-5598. doi: 10.1097/HJH.0000000000001952. URL https://pubmed.ncbi.nlm.nih.gov/30817445/.

[64] Sara S. Al Disi, M. Akhtar Anwar, and Ali H. Eid. Anti-hypertensive herbs and their mechanisms of action: Part i. *Frontiers in Pharmacology*, 6:323, 1 2015. ISSN 16639812. doi: 10.3389/FPHAR.2015.00323. URL /pmc/articles/PMC4717468//pmc/articles/PMC4717468/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4717468/.

[65] M. Akhtar Anwar, Sara S. Al Disi, and Ali H. Eid. Anti-hypertensive herbs and their mechanisms of action: Part ii. *Frontiers in Pharmacology*, 7:50, 2016. ISSN 16639812. doi: 10.3389/FPHAR.2016.00050. URL /pmc/articles/PMC4782109//pmc/articles/PMC4782109/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4782109/.

[66] Ha Nguyen, Olaide A. Odelola, Janani Rangaswami, and Aman Amanullah. A review of nutritional factors in hypertension management. *International Journal of Hypertension*, 2013:12, 2013. ISSN 20900384. doi: 10.1155/2013/698940. URL /pmc/articles/PMC3649175//pmc/articles/PMC3649175/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3649175/.

[67] Luciana Tavares Toscano, Cássia Surama Oliveira da Silva, Lydiane Tavares Toscano, Antônio Eduardo Monteiro de Almeida, Amilton da Cruz Santos, and Alexandre Sérgio Silva. Chia flour supplementation reduces blood pressure in hypertensive subjects. *Plant foods for human nutrition (Dordrecht, Netherlands)*, 69:392–398, 12 2014. ISSN 1573-9104. doi: 10.1007/S11130-014-0452-7. URL https://pubmed.ncbi.nlm.nih.gov/25403867/.

[68] Saman Khalesi, Christopher Irwin, and Matt Schubert. Flaxseed consumption may reduce blood pressure: a systematic review and meta-analysis of controlled trials. *The Journal of nutrition*, 145:758–765, 2015. ISSN 1541-6100. doi: 10.3945/JN.114.205302. URL https://pubmed.ncbi.nlm.nih.gov/25740909/.

[69] Lorna C. Gee and Amrita Ahluwalia. Dietary nitrate lowers blood pressure: Epidemiological, pre-clinical experimental and clinical trial evidence. *Current Hypertension Reports*, 18:1–14, 1 2016. ISSN 15343111. doi: 10.1007/S11906-015-0623-4. URL /pmc/articles/PMC4729801//pmc/articles/PMC4729801/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4729801/.

[70] Elena Jovanovski, Laura Bosco, Kashif Khan, Fei Au-Yeung, Hoang Ho, Andreea Zurbau, Alexandra L. Jenkins, and Vladimir Vuksan. Effect of spinach, a high dietary nitrate source, on arterial stiffness and related hemodynamic

measures: A randomized, controlled trial in healthy adults. *Clinical Nutrition Research*, 4:160, 2015. ISSN 2287-3732. doi: 10.7762/CNR.2015.4.3.160. URL /pmc/articles/PMC4525132//pmc/articles/PMC4525132/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4525132/.

[71] Fooddata central. URL https://fdc.nal.usda.gov/fdc-app.html#/food-details/170683/nutrients.

[72] J. Eales, I. Lenoir-Wijnkoop, S. King, H. Wood, F. J. Kok, R. Shamir, A. Prentice, M. Edwards, J. Glanville, and R. L. Atkinson. Is consuming yoghurt associated with weight management outcomes? results from a systematic review. *International Journal of Obesity (2005)*, 40:731, 5 2016. ISSN 14765497. doi: 10.1038/IJO.2015.202. URL /pmc/articles/PMC4856732//pmc/articles/PMC4856732/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4856732/.

[73] Earl S. Ford and Ali H. Mokdad. Fruit and vegetable consumption and diabetes mellitus incidence among u.s. adults. *Preventive Medicine*, 32:33–39, 1 2001. ISSN 0091-7435. doi: 10.1006/PMED.2000.0772.

[74] Ken Fujioka, Frank Greenway, Judy Sheard, and Yu Ying. The effects of grapefruit on weight and insulin resistance: relationship to the metabolic syndrome. *Journal of medicinal food*, 9:49–54, 3 2006. ISSN 1096-620X. doi: 10.1089/JMF.2006.9.49. URL https://pubmed.ncbi.nlm.nih.gov/16579728/.

[75] Eleanor J. Beck, Susan M. Tosh, Marijka J. Batterham, Linda C. Tapsell, and Xu Feng Huang. Oat beta-glucan increases postprandial cholecystokinin levels, decreases insulin response and extends subjective satiety in overweight subjects. *Molecular nutrition & food research*, 53:1343–1351, 10 2009. ISSN 1613-4133. doi: 10.1002/MNFR.200800343. URL https://pubmed.ncbi.nlm.nih.gov/19753601/.

[76] Richard D. Mattes, Penny M. Kris-Etherton, and Gary D. Foster. Impact of peanuts and tree nuts on body weight and healthy weight loss in adults. *The Journal of nutrition*, 138, 2008. ISSN 1541-6100. doi: 10.1093/JN/138.9.1741S. URL https://pubmed.ncbi.nlm.nih.gov/18716179/.

[77] Elin Östman, Y. Granfeldt, L. Persson, and I. Björck. Vinegar supplementation lowers glucose and insulin responses and increases satiety after a bread meal in healthy subjects. *European journal of clinical nutrition*, 59:983–988, 9 2005. ISSN 0954-3007. doi: 10.1038/SJ.EJCN.1602197. URL https://pubmed.ncbi.nlm.nih.gov/16015276/.

[78] Nuray Z. Unlu, Torsten Bohn, Steven K. Clinton, and Steven J. Schwartz. Carotenoid absorption from salad and salsa by humans is enhanced by the addition of avocado

or avocado oil. *The Journal of nutrition*, 135:431–436, 2005. ISSN 0022-3166. doi: 10.1093/JN/135.3.431. URL https://pubmed.ncbi.nlm.nih.gov/15735074/.

[79] Michael B. Zemel, Warren Thompson, Anita Milstead, Kristin Morris, and Peter Campbell. Calcium and dairy acceleration of weight and fat loss during energy restriction in obese adults. *Obesity research*, 12:582–590, 2004. ISSN 1071-7323. doi: 10.1038/OBY.2004.67. URL https://pubmed.ncbi.nlm.nih.gov/15090625/.

[80] Julie E. Flood and Barbara J. Rolls. Soup preloads in a variety of forms reduce meal energy intake. *Appetite*, 49:626, 11 2007. ISSN 01956663. doi: 10.1016/J.APPET.2007.04.002. URL /pmc/articles/PMC2128765/https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2128765/.

[81] Susanna H.A. Holt, Jennie C. Brand-Miller, and Paul A. Stitt. The effects of equal-energy portions of different breads on blood glucose levels, feelings of fullness and subsequent food intake. *Journal of the American Dietetic Association*, 101:767–773, 2001. ISSN 00028223. doi: 10.1016/S0002-8223(01)00192-4.

[82] Heather J. Leidy, Minghua Tang, Cheryl L.H. Armstrong, Carmen B. Martin, and Wayne W. Campbell. The effects of consuming frequent, higher protein meals on appetite and satiety during weight loss in overweight/obese men. *Obesity (Silver Spring, Md.)*, 19:818–824, 4 2011. ISSN 1930-739X. doi: 10.1038/OBY.2010.203. URL https://pubmed.ncbi.nlm.nih.gov/20847729/.

[83] Margriet A.B. Veldhorst, Margriet S. Westerterp-Plantenga, and Klaas R. Westerterp. Gluconeogenesis and energy expenditure after a high-protein, carbohydrate-free diet. *The American journal of clinical nutrition*, 90:519–526, 9 2009. ISSN 1938-3207. doi: 10.3945/AJCN.2009.27834. URL https://pubmed.ncbi.nlm.nih.gov/19640952/.

[84] Margriet S. Westerterp-Plantenga. Protein intake and energy balance. *Regulatory peptides*, 149:67–69, 8 2008. ISSN 0167-0115. doi: 10.1016/J.REGPEP.2007.08.026. URL https://pubmed.ncbi.nlm.nih.gov/18448177/.

[85] Umesh Kapil. Health consequences of iodine deficiency. *Sultan Qaboos University Medical Journal*, 7:267, 12 2007. ISSN 20750528. URL /pmc/articles/PMC3074887//pmc/articles/PMC3074887/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3074887/.

[86] Julia A. Ello-Martin, Liane S. Roe, Jenny H. Ledikwe, Amanda M. Beach, and Barbara J. Rolls. Dietary energy density in the treatment of obesity: a year-long trial comparing 2 weight-loss diets. *The American journal of clinical nutrition*, 85:1465–1477, 6 2007. ISSN 0002-9165. doi: 10.1093/AJCN/85.6.1465. URL https://pubmed.ncbi.nlm.nih.gov/17556681/.

[87] J. S.Vander Wal, A. Gupta, P. Khosla, and N. V. Dhurandhar. Egg breakfast enhances weight loss. *International journal of obesity (2005)*, 32:1545–1551, 10 2008. ISSN 1476-5497. doi: 10.1038/IJO.2008.130. URL https://pubmed.ncbi.nlm.nih.gov/18679412/.

[88] Jillon S. Vander Wal, Pramod Khosla, K. L.Catherine Jen, Nikhil V. Dhurandhar, Jorene M. Marth, and Nikhil V. Dhurandhar. Short-term effect of eggs on satiety in overweight and obese subjects. *Journal of the American College of Nutrition*, 24: 510–515, 12 2005. ISSN 0731-5724. doi: 10.1080/07315724.2005.10719497. URL https://pubmed.ncbi.nlm.nih.gov/16373948/.

[89] Juan Wang, Xiuming Zhang, Haili Lan, and Weijia Wang. Effect of garlic supplement in the management of type 2 diabetes mellitus (t2dm): a meta-analysis of randomized controlled trials. *Food & Nutrition Research*, 61, 2017. ISSN 16546628. doi: 10.1080/16546628.2017.1377571. URL /pmc/articles/PMC5642189//pmc/articles/PMC5642189/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5642189/.

[90] Annika S. Axelsson, Emily Tubbs, Brig Mecham, Shaji Chacko, Hannah A. Nenonen, Yunzhao Tang, Jed W. Fahey, Jonathan M.J. Derry, Claes B. Wollheim, Nils Wierup, Morey W. Haymond, Stephen H. Friend, Hindrik Mulder, and Anders H. Rosengren. Sulforaphane reduces hepatic glucose production and improves glucose control in patients with type 2 diabetes. *Science Translational Medicine*, 9, 6 2017. ISSN 19466242. doi: 10.1126/SCITRANSLMED.AAH4477/SUPPL_FILE/AAH4477_TABLES_S2_AND_S3.ZIP. URL https://www.science.org/doi/abs/10.1126/scitranslmed.aah4477.

[91] Gang Liu, Marta Guasch-Ferré, Yang Hu, Yanping Li, Frank B. Hu, Eric B. Rimm, Jo Ann E. Manson, Kathryn M. Rexrode, and Qi Sun. Nut consumption in relation to cardiovascular disease incidence and mortality among patients with diabetes mellitus. *Circulation Research*, 124:920–929, 3 2019. ISSN 15244571. doi: 10.1161/CIRCRESAHA.118.314316. URL https://www.ahajournals.org/doi/abs/10.1161/CIRCRESAHA.118.314316.

[92] Yoona Kim, Jennifer B. Keogh, and Peter M. Clifton. Benefits of nut consumption on insulin resistance and cardiovascular risk factors: Multiple potential mechanisms of actions. *Nutrients*, 9, 11 2017. ISSN 20726643. doi: 10.3390/NU9111271. URL /pmc/articles/PMC5707743//pmc/articles/PMC5707743/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5707743/.

[93] V. Vuksan, A. L. Jenkins, C. Brissette, L. Choleva, E. Jovanovski, A. L. Gibbs, R. P. Bazinet, F. Au-Yeung, A. Zurbau, H. V.T. Ho, L. Duvnjak, J. L. Sievenpiper, R. G. Josse, and A. Hanna. Salba-chia (salvia hispanica l.) in the treatment of overweight and obese

patients with type 2 diabetes: A double-blind randomized controlled trial. *Nutrition, metabolism, and cardiovascular diseases : NMCD*, 27:138–146, 2 2017. ISSN 1590-3729. doi: 10.1016/J.NUMECD.2016.11.124. URL https://pubmed.ncbi.nlm.nih.gov/28089080/.

[94] Renée Wilson, Jinny Willis, Richard Gearry, Paula Skidmore, Elizabeth Fleming, Chris Frampton, and Anitra Carr. Inadequate vitamin c status in prediabetes and type 2 diabetes mellitus: Associations with glycaemic control, obesity, and smoking. *Nutrients*, 9, 9 2017. ISSN 20726643. doi: 10.3390/NU9090997. URL /pmc/articles/PMC5622757//pmc/articles/PMC5622757/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5622757/.

[95] Ashish Chaddha and Kim A. Eagle. Omega-3 fatty acids and heart health. *Circulation*, 132:e350–e352, 12 2015. ISSN 15244539. doi: 10.1161/CIRCULATIONAHA.114.015176. URL https://www.ahajournals.org/doi/abs/10.1161/CIRCULATIONAHA.114.015176.

[96] Sylvie S L Leung Yinko, Ken D Stark, George Thanassoulis, and Louise Pilote. Fish consumption and acute coronary syndrome: A meta-analysis. *The American Journal of Medicine*, 127:848–857, 2014. doi: 10.1016/j.amjmed.2014.04.016. URL http://dx.doi.org/10.1016/j.amjmed.2014.04.016.

[97] Renata Micha, Sarah K. Wallace, and Dariush Mozaffarian. Red and processed meat consumption and risk of incident coronary heart disease, stroke, and diabetes: A systematic review and meta-analysis. *Circulation*, 121:2271, 6 2010. ISSN 00097322. doi: 10.1161/CIRCULATIONAHA.109.924977. URL /pmc/articles/PMC2885952//pmc/articles/PMC2885952/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC2885952/.

[98] Shirin Pourafshar, Neda S. Akhavan, Kelli S. George, Elizabeth M. Foley, Sarah A. Johnson, Behnam Keshavarz, Negin Navaei, Anis Davoudi, Elizabeth A. Clark, and Bahram H. Arjmandi. Egg consumption may improve factors associated with glycemic control and insulin sensitivity in adults with pre- and type ii diabetes. *Food & function*, 9:4469–4479, 8 2018. ISSN 2042-650X. doi: 10.1039/C8FO00194D. URL https://pubmed.ncbi.nlm.nih.gov/30073224/.

[99] S. H. Liu, H. I. Chu, T. F. yu, H. C. Hsu, and T. Y. Cheng. Anemia in vitamin c deficiency and its response to iron:. *https://doi.org/10.3181/00379727-46-12078*, 46:603–606, 11 2016. ISSN 15353699. doi: 10.3181/00379727-46-12078. URL https://journals.sagepub.com/doi/abs/10.3181/00379727-46-12078?journalCode=ebma.

[100] Iron deficiency anemia - symptoms and causes - mayo clinic. URL https://www.mayoclinic.org/diseases-conditions/iron-deficiency-anemia/symptoms-causes/syc-20355034.

[101] Thomas G. DeLoughery. Iron deficiency anemia. *Medical Clinics of North America*, 101: 319–332, 3 2017. ISSN 15579859. doi: 10.1016/J.MCNA.2016.09.004.

[102] Iron-deficiency anemia — nhlbi, nih. URL https://www.nhlbi.nih.gov/health-topics/iron-deficiency-anemia.

[103] Iron-deficiency anemia - hematology.org. URL https://www.hematology.org/education/patients/anemia/iron-deficiency.